

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2004, 08:42:16 ; Search time 1556 Seconds
(Without alignments) 557.108 Million cell updates/sec

Title: US-10-006-191-48

Perfect score: 20

Sequence: 1 aaacatgtaactttgtgtca 20

Scoring table: IDENTITY NUC
Gapop.10.0, Gapexc 1.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 1603530

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_hcg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_scs:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_pa:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_scs:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_hcg_hum:*
- 31: em_hcg_inv:*
- 32: em_hcg_other:*
- 33: em_hcg_mus:*
- 34: em_hcg_pin:*
- 35: em_hcg_rtd:*
- 36: em_hcg_mam:*
- 37: em_hcg_vrt:*
- 38: em_sv:*
- 39: em_hcgo_hum:*
- 40: em_hcgo_mus:*
- 41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|----------|---------------------|
| C 1 | 14.8 | 74.0 | 47 | 6 | AR290734 | AR290734 Sequence |
| C 2 | 13.8 | 69.0 | 43 | 6 | AR411025 | AR411025 Sequence |
| C 3 | 13.6 | 68.0 | 25 | 6 | AX805034 | AX805034 Sequence |
| C 4 | 13.6 | 68.0 | 33 | 6 | AR020722 | AR020722 Sequence |
| C 5 | 13.6 | 68.0 | 35 | 6 | AR020629 | AR020629 Sequence |
| C 6 | 13.6 | 68.0 | 35 | 6 | AR160260 | AR160260 Sequence |
| C 7 | 13.6 | 68.0 | 43 | 6 | AX848560 | AX848560 Sequence |
| C 8 | 13.4 | 67.0 | 24 | 6 | 106920 | 106920 Sequence 10 |
| C 9 | 13.4 | 67.0 | 40 | 6 | AR279512 | AR279512 Sequence |
| C 10 | 13.4 | 67.0 | 40 | 6 | AX025258 | AX025258 Sequence |
| C 11 | 13.2 | 66.0 | 22 | 6 | BD089416 | BD089416 A method |
| C 12 | 13.2 | 66.0 | 22 | 6 | BD089578 | BD089578 A method |
| C 13 | 13.2 | 66.0 | 22 | 12 | AB068139 | AB068139 Synthetic |
| C 14 | 13.2 | 66.0 | 22 | 12 | AB068145 | AB068145 Synthetic |
| C 15 | 13.2 | 66.0 | 37 | 8 | AB589559 | AB589559 Arabidops |
| C 16 | 13.2 | 66.0 | 50 | 6 | AX164881 | AX164881 Sequence |
| C 17 | 13.2 | 65.0 | 21 | 6 | AX096014 | AX096014 Sequence |
| C 18 | 12.8 | 64.0 | 22 | 6 | AX511748 | AX511748 Sequence |
| C 19 | 12.8 | 64.0 | 24 | 6 | 136279 | 136279 Sequence 9 |
| C 20 | 12.8 | 64.0 | 24 | 6 | AX119490 | AX119490 Sequence |
| C 21 | 12.6 | 63.0 | 27 | 6 | AR060293 | AR060293 Sequence |
| C 22 | 12.6 | 63.0 | 27 | 6 | E31816 | E31816 Method for |
| C 23 | 12.6 | 63.0 | 27 | 6 | 120953 | 120953 Sequence 20 |
| C 24 | 12.6 | 63.0 | 27 | 6 | AX138047 | AX138047 Sequence |
| C 25 | 12.6 | 63.0 | 37 | 6 | AX687189 | AX687189 Sequence |
| C 26 | 12.6 | 63.0 | 37 | 6 | AX687199 | AX687199 Sequence |
| C 27 | 12.6 | 63.0 | 37 | 6 | AX687219 | AX687219 Sequence |
| C 28 | 12.6 | 63.0 | 39 | 6 | BD160853 | BD160853 Oxidoredu |
| C 29 | 12.6 | 63.0 | 47 | 6 | AR288677 | AR288677 Sequence |
| C 30 | 12.4 | 62.0 | 38 | 6 | AR099782 | AR099782 Sequence |
| C 31 | 12.4 | 62.0 | 41 | 6 | AR099776 | AR099776 Sequence |
| C 32 | 12.4 | 62.0 | 44 | 6 | AR099784 | AR099784 Sequence |
| C 33 | 12.4 | 62.0 | 47 | 6 | AR289287 | AR289287 Sequence |
| C 34 | 12.4 | 62.0 | 50 | 14 | S44878 | S44878 (5'end of A |
| C 35 | 12.4 | 62.0 | 17 | 6 | AX530632 | AX530632 Sequence |
| C 36 | 12.2 | 61.0 | 17 | 6 | AX735739 | AX735739 Sequence |
| C 37 | 12.2 | 61.0 | 20 | 6 | AX645126 | AX645126 Sequence |
| C 38 | 12.2 | 61.0 | 24 | 6 | BD131502 | BD131502 Heat shock |
| C 39 | 12.2 | 61.0 | 25 | 6 | AX532684 | AX532684 Sequence |
| C 40 | 12.2 | 61.0 | 25 | 6 | AX532685 | AX532685 Sequence |
| C 41 | 12.2 | 61.0 | 25 | 6 | AX532686 | AX532686 Sequence |
| C 42 | 12.2 | 61.0 | 25 | 6 | AX532687 | AX532687 Sequence |
| C 43 | 12.2 | 61.0 | 25 | 6 | AX532688 | AX532688 Sequence |
| C 44 | 12.2 | 61.0 | 25 | 6 | AX532689 | AX532689 Sequence |
| C 45 | 12.2 | 61.0 | 25 | 6 | AX532690 | AX532690 Sequence |

ALIGNMENTS

RESULT 1
AR290734/c
LOCUS AR290734 47 bp DNA
DEFINITION Sequence 2469 from patent US 6537751.
ACCESSION AR290734
VERSION AR290734.1 GI:31678018
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
1 (bases 1 to 47)
REFERENCE Cohen,D., Chumakov,I. and Blumenfeld,M.
AUTHORS Biallelic markers for use in constructing a high density
TITLE disequilibrium map of the human genome
JOURNAL Patent: US 6537751-A 2469 25-MAR-2003;

FEATURES

source

Location/Qualifiers
1..47
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match

Best Local Similarity 74.0%; Score 14.8; DB 6; Length 47;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

CY

1 AACATGTAACCTTTGGTCA 20
|||||
33 AACATGTAACCTTTGGTCA 14

Db

RESULT 2

AR411025

LOCUS Sequence 14 from patent US 6635475. 43 bp DNA linear PAT 18-DEC-2003

DEFINITION AR411025

ACCESSION AR411025.1 GI:40162603

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Query Match

Best Local Similarity 69.0%; Score 13.8; DB 6; Length 43;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY

1 AACATGTAACCTTTGGT 17
|||||
1 AACATGTAACCTTTGGT 17

Db

RESULT 3

AX805034/c

LOCUS Sequence 1202 from Patent WO03060160. 25 bp DNA linear PAT 25-NOV-2003

DEFINITION AX805034

ACCESSION AX805034.1 GI:36522175

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Query Match

Best Local Similarity 68.0%; Score 13.6; DB 6; Length 25;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

CY

1 AACATGTAACCTTTGGTCA 20
|||||
20 AACATGTAACCTTTGGTCA 1

Db

RESULT 4

AR020722

LOCUS Sequence 11 from patent US 5789194. 33 bp DNA linear PAT 05-DEC-1998

DEFINITION AR020722

ACCESSION AR020722.1 GI:3975337

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Query Match

Best Local Similarity 68.0%; Score 13.6; DB 6; Length 33;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

CY

1 AACATGTAACCTTTGGTCA 20
|||||
13 AACATGTAACCTTTGGTCA 32

Db

RESULT 5

AR020629/c

LOCUS Sequence 15 from patent US 5789184. 35 bp DNA linear PAT 05-DEC-1998

DEFINITION AR020629

ACCESSION AR020629.1 GI:3975244

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Query Match

Best Local Similarity 68.0%; Score 13.6; DB 6; Length 35;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

CY

1 AACATGTAACCTTTGGTCA 20
|||||
20 AACATGTAACCTTTGGTCA 1

Db

RESULT 6

AR160260/c

LOCUS Sequence 6 from patent US 6255059. 35 bp DNA linear PAT 17-OCT-2001

DEFINITION AR160260

ACCESSION AR160260.1 GI:16223904

VERSION

KEYWORDS

```

SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE    1 (bases 1 to 35)
AUTHORS      Klein,C.A., Murphy,A.J.M., Fowlkes,D.M., Broach,J., Manfredi,J.,
              Paul,J. and Truheart,J.
TITLE        Methods for identifying G protein coupled receptor effectors
JOURNAL      Patent: US 625059-A 6 03-0UL-2001;
FEATURES     Location/Qualifiers
              source
                1..35
                /organism="Unknown"
                /mol_type="unassigned DNA"

ORIGIN
Query Match      68.0%; Score 13.6; DB 6; Length 35;
Best Local Similarity 80.0%; Pred. No. 4.4e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Cy 1 AAACATGTAACCTTTGGTCA 20
    ||| ||| ||| ||| |||
Db 20 AAACATGTAATTTGATCA 1

RESULT 7
AX484560      43 bp  DNA      linear  PAT 16-AUG-2002
LOCUS         Sequence 1860 from Patent WO02053728.
ACCESSION    AX484560
VERSION      AX484560.1 GI:22318912
KEYWORDS
SOURCE
ORGANISM     Candida albicans
              Candida albicans
              Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
              Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE
AUTHORS      Roemer,T., Jiaerg,B., Boone,C., Bussey,H. and Ohlsen,K.L.
TITLE        Gene disruption methodologies for drug target discovery
JOURNAL      Patent: WO 02053728-A 1860 11-JUL-2002;
              Elittra Pharmaceuticals, Inc. (US)
FEATURES     Location/Qualifiers
              source
                1..43
                /organism="Candida albicans"
                /mol_type="unassigned DNA"
                /db_xref="taxon:5476"

ORIGIN
Query Match      68.0%; Score 13.6; DB 6; Length 43;
Best Local Similarity 80.0%; Pred. No. 4.4e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Cy 1 AAACATGTAACCTTTGGTCA 20
    ||| ||| ||| ||| |||
Db 11 AAACATTTAGTTTGGTAA 30

RESULT 8
LOCUS         106920      24 bp  DNA      linear  PAT 02-DEC-1994
DEFINITION    Sequence 10 from Patent EP 0340948.
ACCESSION    106920
VERSION      106920.1 GI:589838
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE    1 (bases 1 to 24)
AUTHORS      Wilcox,E., Edwards,D.L., Schwab,G.E., Thompson,M. and Culver,P.
TITLE        Novel hybrid pesticidal toxins
JOURNAL      Patent: EP 0340948-A1 10 08-NOV-1989;
FEATURES     Location/Qualifiers
              source
                1..24
                /organism="Unknown"
                /mol_type="unassigned DNA"

ORIGIN
Query Match      67.0%; Score 13.4; DB 6; Length 24;
Best Local Similarity 93.3%; Pred. No. 5.6e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 5 ATCTACTTTTGGTC 19
    ||| ||| ||| ||| |||
Db 8 ATGTACTTTTGGTC 22

RESULT 9
LOCUS         AR279512      40 bp  DNA      linear  PAT 10-APR-2003
DEFINITION    Sequence 41 from patent US 6514706.
ACCESSION    AR279512
VERSION      AR279512.1 GI:29714332
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE    Unclassified.
AUTHORS      Von Kalle,C. and Schmidt,M.
TITLE        Linear amplification mediated PCR (LAM PCR)
JOURNAL      Patent: US 6514706-A 41 04-FEB-2003;
FEATURES     Location/Qualifiers
              source
                1..40
                /organism="Unknown"
                /mol_type="genomic DNA"

ORIGIN
Query Match      67.0%; Score 13.4; DB 6; Length 40;
Best Local Similarity 87.5%; Pred. No. 5.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 2 AACATGTAACCTTTGG 17
    ||| ||| ||| ||| |||
Db 14 AACATGTAACCTGNTGG 29

RESULT 10
LOCUS         AX025258      40 bp  DNA      linear  PAT 16-SEP-2000
DEFINITION    Sequence 41 from Patent DE19849348.
ACCESSION    AX025258
VERSION      AX025258.1 GI:10186972
KEYWORDS
SOURCE      Macaca mulatta (rhesus monkey)
              Macaca mulatta
              Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
              Cercopitheciinae; Macaca.
REFERENCE    1
AUTHORS      Schmidt,M. and Kalle,C.V.
JOURNAL      Patent: DE 19849348-A 41 27-APR-2000;
              UNIV LUDWIGS ALBERT (DE)
FEATURES     Location/Qualifiers
              source
                1..40
                /organism="Macaca mulatta"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9544"
                /note="Die ersten 10 Nucleotide von pLN 5'LTR
                (NCBI/Genbank: M28245)"

ORIGIN
Query Match      67.0%; Score 13.4; DB 6; Length 40;
Best Local Similarity 87.5%; Pred. No. 5.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 2 AACATGTAACCTTTGG 17
    ||| ||| ||| ||| |||
Db 14 AACATGTAACCTGNTGG 29

```

RESULT 11
BD089416/c
LOCUS BD089416 22 bp DNA linear PAT 27-AUG-2002
DEFINITION A method of arraying genome clone.
ACCESSION BD089416
VERSION BD089416.1 GI:22635026
KEYWORDS JP 2001321190-A/1660.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
1 (bases 1 to 22)
REFERENCE Soeda,E.
AUTHORS A method of arraying genome clone
TITLE Patent: JP 2001321190-A 1660 20-NOV-2001;
JOURNAL THE INSTITUTE OF PHYSICAL AND CHEMICAL RESEARCH, YUGENKAISHA
GENOTECs

COMMENT OS Artificial Sequence
PN UP 2001321190-A/1660
PD 20-NOV-2001
PF 12-MAR-2001 JP 2001068285
PI EIICHI SOEDA
PC C12N15/09,C12N15/00,C12Q1/68,G01N33/53,G01N33/566, PC
C12N15/00
PC C12N15/00
CC Description of Artificial Sequence:Synthetic DNA FH Key
FT source Location/Qualifiers
1..22
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 66.0%; Score 13.2; DB 6; Length 22;
Best Local Similarity 83.3%; Pred. No. 7.1e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 ACATGTAACCTTTGGTCA 20
19 ACATGTAACATGTGGCCA 2

Db

RESULT 12
BD089578/c
LOCUS BD089578 22 bp DNA linear PAT 27-AUG-2002
DEFINITION A method of arraying genome clone.
ACCESSION BD089578
VERSION BD089578.1 GI:22635188
KEYWORDS JP 2001321190-A/1822.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
1 (bases 1 to 22)
REFERENCE Soeda,E.
AUTHORS A method of arraying genome clone
TITLE Patent: JP 2001321190-A 1822 20-NOV-2001;
JOURNAL THE INSTITUTE OF PHYSICAL AND CHEMICAL RESEARCH, YUGENKAISHA
GENOTECs

COMMENT OS Artificial Sequence
PN UP 2001321190-A/1822
PD 20-NOV-2001
PF 12-MAR-2001 JP 2001068285
PI EIICHI SOEDA
PC C12N15/09,C12N15/00,C12Q1/68,G01N33/53,G01N33/566, PC
C12N15/00
PC C12N15/00
CC Description of Artificial Sequence:Synthetic DNA FH Key
FT source Location/Qualifiers
1..22

FEATURES FT Location/Qualifiers
source 1..22
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 66.0%; Score 13.2; DB 6; Length 22;
Best Local Similarity 83.3%; Pred. No. 7.1e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 ACATGTAACCTTTGGTCA 20
19 ACATGTAACATGTGGCCA 2

Db

RESULT 13
AB068139/c
LOCUS AB068139 22 bp DNA linear SYN 21-MAY-2003
DEFINITION Synthetic construct DNA, reverse primer for human STS
ACCESSION AB068139
VERSION AB068139.1 GI:15128943
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
1
REFERENCE Chen,Y.Z., Hayashii,Y., Wu,J.G., Takaoka,E., Maekawa,K.,
AUTHORS Watanabe,N., Inazawa,J., Hosoda,F., Arai,Y., Mizushima,H.,
Mochizuki,A., Ohira,M., Nakagawara,A., Liu,S., Hoshi,M., Horii,A.
and Soeda,E.
TITLE A BAC-based STS-content map spanning a 35-Mb region of human
JOURNAL Chromosome 1p35-p36
MEDLINE Genomics 74 (1), 55-70 (2001)
PUBMED 21269192
REFERENCE 2 (bases 1 to 22)
AUTHORS Horii,A.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-2001) Akira Horii, Tohoku University School of
Medicine, Molecular Pathology; 2-1 Seiryomachi, Aoba-ku, Sendai,
Miyagi 980-8575, Japan (E-mail: horii@mail.cc.tohoku.ac.jp,
Tel:81-22-717-8042, Fax:81-22-717-8047)
LOCATION/Qualifiers
1..22
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

misc_feature
1..22
/note="reverse primer for human STS sts-stcDB443043 at
1p36
sts-stcDB443043 obtained from clones B56A11, B239M20,
B215H8, B239P22, B239P22, B88A11, Human BAC library
RPC1-11"

ORIGIN
Query Match 66.0%; Score 13.2; DB 12; Length 22;
Best Local Similarity 83.3%; Pred. No. 7.1e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 ACATGTAACCTTTGGTCA 20
19 ACATGTAACATGTGGCCA 2

Db

RESULT 14
AB068145/c
LOCUS AB068145 22 bp DNA linear SYN 21-MAY-2003
DEFINITION Synthetic construct DNA, reverse primer for human STS
ACCESSION AB068145

```

VERSION      AB068145.1  GI:15128949
KEYWORDS
SOURCE       synthetic construct
ORGANISM     synthetic construct
REFERENCE    1
AUTHORS      Chen, Y.Z., Hayashi, Y., Wu, J.G., Takaoka, E., Maekawa, K.,
              Matanabe, N., Inazawa, J., Hosoda, F., Arai, Y., Mizushima, H.,
              Morohashi, A., Ohira, M., Nakagawara, A., Liu, S., Hoshi, M., Horii, A.
              and Soda, E.
TITLE        A BAC-based STS-content map spanning a 35-Mb region of human
              chromosome 1p35-p36
JOURNAL      Genomics 74 (1), 55-70 (2001)
MEDLINE      21269192
PUBMED       11374902
REFERENCE    2 (bases 1 to 22)
AUTHORS      Horii, A.
TITLE        Direct Submission
              Submitted (04-AUG-2001) Akira Horii, Tohoku University School of
              Medicine, Molecular Pathology, 2-1 Seiryomachi, Aoba-ku, Sendai,
              Miyagi 980-8575, Japan (E-mail: horii@mail.cc.tohoku.ac.jp,
              Tel:81-22-717-8042, Fax:81-22-717-8047)
FEATURES
  source
    1..22
    location/Qualifiers
      /organism="synthetic construct"
      /mol_type="genomic DNA"
      /db_xref="taxon:32630"
  misc_feature
    1..22
    /note="reverse primer for human STS sts-A002035 at 1p36
    sts-A002035 obtained from clones B262K21, B239W20, B215H8,
    B239P22, B239P22, B31010.6, B262K21, Human BAC library
    RPC1-11"

ORIGIN
Query Match      66.0%; Score 13.2; DB 12; Length 22;
Best Local Similarity 83.3%; Pred. No. 7.1e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      3 ACATGTAACATTTGGTCA 20
        |||||
        19 ACATGTAACATGTGCCA 2

RESULT 15
AJ589959/c      37 bp DNA linear PLN 23-OCT-2003
LOCUS          Arabidopsis thaliana T-DNA flanking sequence, left border, clone
561E05.
ACCESSION      AJ589959
KEYWORDS       left border; T-DNA flanking sequence.
SOURCE         Arabidopsis thaliana (chale cress)
ORGANISM       Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE      1
AUTHORS        Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F.,
              Chauvin, S., Bechtold, N., Crund, C., DeRose, R., Pelletier, G.,
              Lepoint, L., Caboche, M., and Lecharny, A.
TITLE          T-DNA integration into the Arabidopsis genome depends on sequences
              of pre-insertion sites
JOURNAL        EMBO Rep. 3 (12), 1152-1157 (2002)
MEDLINE        22363535
PUBMED         12446565
REFERENCE      2 (bases 1 to 37)
AUTHORS        Balzerque, S.
TITLE          Direct Submission
              Submitted (23-OCT-2003) Balzerque S., UMRGV, INRA/CNRS, 2 rue
              Gaston Cremieux, 91057 Evry Cedex, FRANCE
              PCR was performed on DNA from transformants of Arabidopsis thaliana
              plants from INRA (Versailles). The DNA fragment(s) resulting from
              the PCR were directly sequenced from the left or the right border

```

to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbsgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).

```

FEATURES
  source
    1..37
    location/Qualifiers
      /organism="Arabidopsis thaliana"
      /mol_type="genomic DNA"
      /cultivar="Wassilewskija"
      /db_xref="taxon:3702"
      /clone="561E05"
      /clone_1nb="Arabidopsis thaliana T-DNA insertion lines"
  misc_feature
    1..37
    /note="T-DNA flanking sequence
    left border"

ORIGIN
Query Match      66.0%; Score 13.2; DB 8; Length 37;
Best Local Similarity 83.3%; Pred. No. 7e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AAACATGTAACATTTGGT 18
        |||||
        19 AAATTGTAAATTTGGT 2

Search completed: March 27, 2004, 10:00:26
Job time : 1558 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2004, 09:23:56 ; Search time 75 Seconds
(without alignments)
147,987 Million cell updates/sec

Title: US-10-006-191-48

Perfect score: 20

Sequence: 1 aaacatgtaactttgtca 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 839752

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued_Patents_NA*

1: /cgn2_6/pdata/2/ina/5A.COMB.seq:*
2: /cgn2_6/pdata/2/ina/5B.COMB.seq:*
3: /cgn2_6/pdata/2/ina/6A.COMB.seq:*
4: /cgn2_6/pdata/2/ina/6B.COMB.seq:*
5: /cgn2_6/pdata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/pdata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 14.8 | 74.0 | 47 | 4 | US-09-422-978-2469 |
| 2 | 13.8 | 69.0 | 43 | 4 | US-09-627-746-14 |
| 3 | 13.6 | 68.0 | 33 | 1 | US-08-450-944-11 |
| 4 | 13.6 | 68.0 | 33 | 5 | PCT-US96-07709-11 |
| 5 | 13.6 | 68.0 | 35 | 1 | US-08-464-531-15 |
| 6 | 13.6 | 68.0 | 35 | 2 | US-08-461-598-15 |
| 7 | 13.6 | 68.0 | 35 | 3 | US-08-322-137-15 |
| 8 | 13.6 | 68.0 | 35 | 3 | US-08-582-333A-6 |
| 9 | 13.4 | 67.0 | 40 | 4 | US-09-830-337-41 |
| 10 | 13.2 | 66.0 | 30 | 4 | US-09-684-579-14 |
| 11 | 13.2 | 66.0 | 31 | 4 | US-09-684-579-16 |
| 12 | 12.8 | 64.0 | 27 | 1 | US-08-411-389-9 |
| 13 | 12.6 | 63.0 | 24 | 1 | US-08-381-280-20 |
| 14 | 12.6 | 63.0 | 27 | 2 | US-08-445-533-20 |
| 15 | 12.6 | 63.0 | 27 | 2 | US-09-052-085-20 |
| 16 | 12.6 | 63.0 | 47 | 4 | US-09-422-978-412 |
| 17 | 12.6 | 63.0 | 38 | 3 | US-09-178-089-9 |
| 18 | 12.4 | 62.0 | 41 | 3 | US-09-178-089-3 |
| 19 | 12.4 | 62.0 | 44 | 3 | US-09-178-089-11 |
| 20 | 12.4 | 62.0 | 47 | 4 | US-09-422-978-1022 |
| 21 | 12.2 | 61.0 | 30 | 3 | US-09-142-078-51 |
| 22 | 12.2 | 61.0 | 30 | 3 | US-09-357-141-51 |
| 23 | 12.2 | 61.0 | 30 | 4 | US-09-533-889-51 |
| 24 | 12.2 | 61.0 | 30 | 4 | US-09-142-080-51 |
| 25 | 12.2 | 61.0 | 31 | 3 | US-09-142-078-59 |
| 26 | 12.2 | 61.0 | 31 | 3 | US-09-357-141-59 |
| 27 | 12.2 | 61.0 | 31 | 4 | US-09-533-889-59 |

| | | | | | | |
|----|------|------|----|---|---------------------|---------------------|
| 28 | 12.2 | 61.0 | 31 | 4 | US-09-142-080-59 | Sequence 59, Appl |
| 29 | 12 | 60.0 | 29 | 3 | US-09-029-819-9 | Sequence 9, Appl |
| 30 | 12 | 60.0 | 39 | 3 | US-08-881-037-87 | Sequence 87, Appl |
| 31 | 12 | 60.0 | 40 | 2 | US-08-857-946-162 | Sequence 162, Appl |
| 32 | 12 | 60.0 | 40 | 3 | US-08-970-740-162 | Sequence 162, Appl |
| 33 | 12 | 60.0 | 40 | 4 | US-09-359-304B-34 | Sequence 34, Appl |
| 34 | 12 | 60.0 | 47 | 4 | US-09-641-638-821 | Sequence 821, Appl |
| 35 | 12 | 60.0 | 47 | 4 | US-09-422-978-2964 | Sequence 2964, Appl |
| 36 | 11.8 | 59.0 | 20 | 4 | US-09-484-617-96 | Sequence 96, Appl |
| 37 | 11.8 | 59.0 | 20 | 4 | US-09-601-144-23 | Sequence 23, Appl |
| 38 | 11.8 | 59.0 | 21 | 4 | US-09-422-978-8370 | Sequence 8370, Appl |
| 39 | 11.8 | 59.0 | 23 | 4 | US-09-747-391-162 | Sequence 162, Appl |
| 40 | 11.8 | 59.0 | 24 | 2 | US-08-713-928B-11 | Sequence 11, Appl |
| 41 | 11.8 | 59.0 | 36 | 1 | US-08-454-097-45 | Sequence 45, Appl |
| 42 | 11.8 | 59.0 | 36 | 4 | US-08-185-359-45 | Sequence 45, Appl |
| 43 | 11.8 | 59.0 | 38 | 3 | US-09-371-772B-7881 | Sequence 7881, Appl |
| 44 | 11.8 | 59.0 | 48 | 3 | US-08-290-995-13 | Sequence 13, Appl |
| 45 | 11.6 | 58.0 | 20 | 3 | US-08-750-141A-12 | Sequence 12, Appl |

ALIGNMENTS

RESULT 1
US-09-422-978-2469/C
Sequence 2469, Application US/09422978
Patent No. 6537751
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marla
APPLICANT: Chumakov, Ilya
TITLE OF INVENTION: Biallelic markers for use in constructing a high density....
FILE REFERENCE: GENST.020CPI
CURRENT APPLICATION NUMBER: US/09/422,978
EARLIER FILING DATE: 1999-10-20
EARLIER APPLICATION NUMBER: US 09/298,850
EARLIER FILING DATE: 1999-04-21
EARLIER APPLICATION NUMBER: US 60/109,732
EARLIER FILING DATE: 1998-11-23
EARLIER APPLICATION NUMBER: US 60/082,614
EARLIER FILING DATE: 1998-04-21
NUMBER OF SEQ ID NOS: 11796
SEQ ID NO 2469
LENGTH: 47
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 24
OTHER INFORMATION: 99-11179-239 : polymorphic base C or T
US-09-422-978-2469

Query Match 74.0%; Score 14.8; DB 4; Length 47;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAACATGTAACCTTTGTCA 20
DB 33 AAACATGTAACCTTTGTCA 14

RESULT 2
US-09-627-746-14
Sequence 14, Application US/09627746
Patent No. 6635475
GENERAL INFORMATION:
APPLICANT: Helmann, John
TITLE OF INVENTION: Bacillus subtilis Extracytoplasmic Function Sigma Factor
FILE REFERENCE: 10845-125
CURRENT APPLICATION NUMBER: US/09/627,746
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/146,466
PRIOR FILING DATE: 1999-07-30

NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patent version 3.0
SEQ ID NO 14
LENGTH: 43
TYPE: DNA
ORGANISM: Bacillus subtilis
US-09-627-746-14

Query Match 69.0%; Score 13.8; DB 4; Length 43;
Best Local Similarity 88.2%; Pred. No. 4.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAACATGTAAGCTTTGG 17
DB 1 AAATATGATGACTTTGGT 17

RESULT 3

US-08-450-944-11
Sequence 11, Application US/08450944
Patent No. 5789194
GENERAL INFORMATION:
APPLICANT: Tripp, Cynthia A.
APPLICANT: Wisniewski, Nancy
TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN
TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln St., Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,944
FILING DATE: 23-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: 2618-30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA primer
US-08-450-944-11

Query Match 68.0%; Score 13.6; DB 1; Length 33;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAACATGTAAGCTTTGGTCA 20
DB 13 AAATATGATGACTTTGGTGA 32

RESULT 4

PCT-US96-07709-11
Sequence 11, Application PC/TUS9607709
GENERAL INFORMATION:
APPLICANT: Tripp, Cynthia A.

APPLICANT: Wisniewski, Nancy
TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN
TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln St., Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/07709
FILING DATE: 23-MAY-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-30-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..33
OTHER INFORMATION: /label= PRIMER
PCT-US96-07709-11

Query Match 68.0%; Score 13.6; DB 5; Length 33;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAACATGTAAGCTTTGGTCA 20
DB 13 AAATATGATGACTTTGGTGA 32

RESULT 5

US-08-464-531-15/C
Sequence 15, Application US/08464531
Patent No. 5789184
GENERAL INFORMATION:
APPLICANT: FOWLES, Dana M.
APPLICANT: BROACH, Jim
APPLICANT: MANFREDI, John
APPLICANT: KLEIN, Christine
APPLICANT: MURPHY, Andrew J.
APPLICANT: PAUL, Jeremy
APPLICANT: TRUEHEART, Joshua
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE
TITLE OF INVENTION: PHENOLIC SYSTEM PROTEIN SUBROGATES, AND USES THEREFOR
NUMBER OF SEQUENCES: 119
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEWMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,531
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/322,137
FILING DATE: 13-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/309,313
FILING DATE: 20-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,328
FILING DATE: 31-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/041,431
FILING DATE: 31-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: FOLWIKES-2G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: synthetic DNA
US-08-464-531-15

Query Match 68.0%; Score 13.6; DB 1; Length 35;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAACATGTACTTTGGTCA 20
DB 20 AAACATGTAAATTGGATCA 1

RESULT 6
US-08-461-598-15/c
Sequence 15, Application US/08461598
Patent No. 5976951
GENERAL INFORMATION:
APPLICANT: FOLWIKES, Dana M.
APPLICANT: BROACH, Jim
APPLICANT: MANFREDI, John
APPLICANT: KLEIN, Christine
APPLICANT: MURPHY, Andrew J.
APPLICANT: PAUL, Jeremy
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE
TITLE OF INVENTION: PHEROMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR
NUMBER OF SEQUENCES: 119
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/461,598
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/322,137
FILING DATE: 13-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/309,313
FILING DATE: 20-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,328
FILING DATE: 31-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/041,431
FILING DATE: 31-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: FOLWIKES-2F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: synthetic DNA
US-08-461-598-15

Query Match 68.0%; Score 13.6; DB 2; Length 35;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAACATGTACTTTGGTCA 20
DB 20 AAACATGTAAATTGGATCA 1

RESULT 7
US-08-322-137-15/c
Sequence 15, Application US/08322137
Patent No. 610042
GENERAL INFORMATION:
APPLICANT: FOLWIKES, Dana M.
APPLICANT: BROACH, Jim
APPLICANT: MANFREDI, John
APPLICANT: KLEIN, Christine
APPLICANT: MURPHY, Andrew J.
APPLICANT: PAUL, Jeremy
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE
TITLE OF INVENTION: PHEROMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR
NUMBER OF SEQUENCES: 119
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,137
FILING DATE: 13-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/309,313
FILING DATE: 20-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,328
FILING DATE: 31-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/041,431
FILING DATE: 31-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: FOLWES-2C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ. ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: synthetic DNA
US-08-322-137-15

Query Match 68.0%; Score 13.6; DB 3; Length 35;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AACATGTACTTTGGTCA 20
DB 20 AACATGTACTTTGGTCA 1

RESULT 8
US-08-582-333a-6/c
Sequence 6, Application US/08582333A
Patent No. 6255059
GENERAL INFORMATION:
APPLICANT: Klein, Christine A.
APPLICANT: Murphy, Andrew J. M.
TITLE OF INVENTION: Methods and Compositions for
IDENTIFYING IDENTIFYING RECEPTOR EFFECTORS
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/582,333A
FILING DATE: 17-JAN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Catherine J. Kara
REGISTRATION NUMBER: 41,106
REFERENCE/DOCKET NUMBER: CPI-012CPS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-4214
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid
US-08-582-333a-6

Query Match 68.0%; Score 13.6; DB 3; Length 35;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AACATGTACTTTGGTCA 20
DB 20 AACATGTACTTTGGTCA 1

RESULT 9
US-09-830-337-41
Sequence 41, Application US/09830337
Patent No. 654706
GENERAL INFORMATION:
APPLICANT: von Kalle, Christof
APPLICANT: Schmidt, Manfred
TITLE OF INVENTION: LINEAR AMPLIFICATION MEDIATED PCR (=LAM PCR)
FILE REFERENCE: 0147-0227P
CURRENT APPLICATION NUMBER: US/09/830,337
CURRENT FILING DATE: 2001-04-26
NUMBER OF SEQ. ID NOS: 50
SOFTWARE: PatentIn Ver. 2.1
SEQ. ID NO. 41
LENGTH: 40
TYPE: DNA
ORGANISM: Macaca mulatta
FEATURE:
NAME/KEY: LTR
LOCATION: (31)..(40)
OTHER INFORMATION: The first 10 nucleotides of pLN 5'LTR
OTHER INFORMATION: (NCBI/Genbank: M28245)
US-09-830-337-41

Query Match 67.0%; Score 13.4; DB 4; Length 40;
Best Local Similarity 87.5%; Pred. No. 6.8e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AACATGTACTTTGG 17
DB 14 AACATGTACTTTGG 29

RESULT 10
US-09-684-579-14/c
Sequence 14, Application US/09684579
Patent No. 6670450
GENERAL INFORMATION:
APPLICANT: Madhwa, Renu
APPLICANT: Kaul, Sunil C.
APPLICANT: Reddel, Roger R.
TITLE OF INVENTION: PROTEIN AND GENE INVOLVED IN MYOCYTE
DIFFERENTIATION
FILE REFERENCE: 06501-066001
CURRENT APPLICATION NUMBER: US/09/684,579
CURRENT FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: PCT/JP99/01913
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: JAPAN 10/115975
PRIOR FILING DATE: 1998-04-10
NUMBER OF SEQ. ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ. ID NO. 14
LENGTH: 30
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: primer for PCR
US-09-684-579-14

Query Match 66.0%; Score 13.2; DB 4; Length 30;
Best Local Similarity 83.3%; Pred. No. 8.2e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 ACATGTAACCTTTGGTCA 20
|||||
DB 28 ACATGAACATTAGGTCA 11

RESULT 11
US-09-684-579-16/c
Sequence 16, Application US/09684579
Patent No. 6670450
GENERAL INFORMATION:
APPLICANT: Madhwa, Renu
APPLICANT: Kaul, Sunil C.
APPLICANT: Reddel, Roger R.
TITLE OF INVENTION: PROTEIN AND GENE INVOLVED IN MYOCYTE
FILE REFERENCE: 06501-066001
CURRENT APPLICATION NUMBER: US/09/684,579
CURRENT FILING DATE: 2000-10-06
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: PCT/J99/01913
PRIOR FILING DATE: 1998-04-10
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 31
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: primer for PCR
US-09-684-579-16

Query Match 66.0%; Score 13.2; DB 4; Length 31;
Best Local Similarity 83.3%; Pred. No. 8.3e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 ACATGTAACCTTTGGTCA 20
|||||
DB 29 ACATGAACATTAGGTCA 12

RESULT 12
US-08-411-389-9
Sequence 9, Application US/08411389
Patent No. 5605799
GENERAL INFORMATION:
APPLICANT: White, Raymond L.
APPLICANT: Cawthon, Richard M.
APPLICANT: Li, Ying
TITLE OF INVENTION: SOMATIC MUTATIONS IN THE
TITLE OF INVENTION: NEUROFIBROMATOSIS TYPE 1 GENE IN HUMAN TUMORS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue NW, Suite 1000
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/411,389
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/047,068

FILING DATE: 16-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 19780-107116
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-411-389-9

Query Match 64.0%; Score 12.8; DB 1; Length 24;
Best Local Similarity 87.5%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AACATGTAACCTTTG 16
|||||
DB 8 AACATGAACCTTTAG 23

RESULT 13
US-08-381-280-20
Sequence 20, Application US/08381280
GENERAL INFORMATION:
APPLICANT: Kazuhiko NISHITANI et al.
TITLE OF INVENTION: ENDO-XYLOGLUCAN TRANSFERASE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/381,280
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/037,281
FILING DATE: March 26, 1993
APPLICATION NUMBER: 07/929,513
FILING DATE: August 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid

```

1      HYPOTHETICAL:
2      ANTI-SENSE:
3      FRAGMENT TYPE:
4      ORIGINAL SOURCE:
5      ORGANISM:
6      STRAIN:
7      INDIVIDUAL ISOLATE:
8      DEVELOPMENTAL STAGE:
9      HAPLOTYPE:
10     TISSUE TYPE:
11     CELL TYPE:
12     CELL LINE:
13     ORGANELLER:
14     IMMEDIATE SOURCE:
15     LIBRARY:
16     CLONE:
17     POSITION IN GENOME:
18     CHROMOSOME/SEGMENT:
19     MAP POSITION:
20     UNITS:
21     FEATURE:
22     NAME/KEY:
23     LOCATION:
24     IDENTIFICATION METHOD:
25     OTHER INFORMATION:
26     PUBLICATION INFORMATION:
27     AUTHORS:
28     TITLE:
29     JOURNAL:
30     VOLUME:
31     ISSUE:
32     PAGES:
33     DATE:
34     DOCUMENT NUMBER:
35     FILING DATE:
36     PUBLICATION DATE:
37     RELEVANT RESIDUES IN SEQ ID NO
38     OS-08-361-280-20

```

| | | | | |
|--------------------------|-------|--------------------|-----------|------------|
| Query Match | 63.0% | Score 12.6; | DB 1; | Length 27; |
| Best Local Similarity | 78.9% | Pred. No. 1.6e+03; | | |
| Matches 15; Conservative | 0; | Mismatches 4; | Indels 0; | Gaps 0; |

| | | | |
|----|---|---------------------------------------|----|
| QY | 1 | A A C A T G T A C T T T T G G T C | 19 |
| | | | |
| Db | 8 | A G A C A T G T A A T T T T A G G C C | 26 |

US RESULT 14
US-08-445-533-20
Sequence 20, Application US/08445533
Patent No. 5840550
GENERAL INFORMATION:
APPLICANT: KAKUHIKO NISHITANI et al.
TITLE OF INVENTION: ENDO-XYLOGUCAN TRANSFERASE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,533
FILING DATE: May 22, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/381,280

FILING DATE: January 31, 1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/037,281
 FILING DATE: March 26, 1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/929,513
 FILING DATE: August 14, 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Warren M. Cheek, Jr.
 REGISTRATION NUMBER: 33,367
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-8850
 TELEFAX:
 TELEX:
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 27 bases
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: other nucleic acid
 HYPOTHEetical:
 ANTI-SENSE:
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 ORGANISM:
 STRAIN:
 INDIVIDUAL ISOLATE:
 DEVELOPMENTAL STAGE:
 HAPLOTYPE:
 TISSUE TYPE:
 CELL TYPE:
 CELL LINE:
 ORGANELLER:
 IMMEDIATE SOURCE:
 LIBRARY:
 CLONE:
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT:
 MAP POSITION:
 UNITS:
 FEATURE:
 NAME/KEY:
 LOCATION:
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 PUBLICATION INFORMATION:
 AUTHORS:
 TITLE:
 JOURNAL:
 VOLUME:
 ISSUE:
 PAGES:
 DATE:
 DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO:

| | | | | |
|--------------------------|-------|-------------------|----------|-----------|
| Query Match | 63.0% | Score 12.6 | DB 2 | Length 27 |
| Best Local Similarity | 78.9% | Pred. No. 1.6e+03 | | |
| Matches 15; Conservative | 0 | Mismatches 4 | Indels 0 | Gaps 0 |

| | | | | |
|----|--|---|----------------------|----|
| OY | | 1 | AACATGTAAGTTTGGTC | 19 |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| D6 | | 8 | AGACATGTAATTCTTAGGCC | 26 |

RESULT 15
US-09-052-085-20
: Sequence 20, Application US/09052085

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2004, 08:16:36 ; Search time 338 Seconds
(without alignments)
251.373 Million cell updates/sec

Title: US-10-006-191-48

Perfect score: 20

Sequence: 1 aaacatgtaacttttggtca 20

Scoring table: IDENTITY NUC

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3185356

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: N Geneseq 29Jan04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002s:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|-----------------------|
| 1 | 20 | 100.0 | 20 | ADB25655 | ADB25655 Human con |
| 2 | 19 | 95.0 | 20 | ADB25671 | ADB25671 Human con |
| 3 | 15.2 | 76.0 | 25 | AC156439 | AC156439 Human mic |
| 4 | 15 | 75.0 | 20 | ADB25654 | ADB25654 Human con |
| 5 | 15 | 75.0 | 25 | AC149802 | AC149802 Human mic |
| 6 | 14.8 | 74.0 | 25 | AB093618 | AB093618 Human DIS |
| 7 | 14.8 | 74.0 | 25 | AC115905 | AC115905 Human mic |
| 8 | 14.2 | 71.0 | 25 | ACK05946 | ACK05946 Human con |
| 9 | 14 | 70.0 | 20 | ADB25670 | ADB25670 Human con |
| 10 | 13.6 | 68.0 | 25 | AC157654 | AC157654 Human mic |
| 11 | 13.6 | 68.0 | 25 | AC117811 | AC117811 Human mic |
| 12 | 13.6 | 68.0 | 25 | AC156438 | AC156438 Human mic |
| 13 | 13.6 | 68.0 | 25 | ADD20567 | ADD20567 Oreochochrom |
| 14 | 13.6 | 68.0 | 26 | AA088537 | AA088537 B. thuring |
| 15 | 13.6 | 68.0 | 33 | AA051400 | AA051400 D. immiti |
| 16 | 13.6 | 68.0 | 35 | AAA88347 | AAA88347 STE2 muta |
| 17 | 13.6 | 68.0 | 35 | ACA61808 | ACA61808 Yeast STE |
| 18 | 13.6 | 68.0 | 43 | AB227913 | AB227913 Candida e |
| 19 | 13.6 | 68.0 | 47 | AA268122 | AA268122 Human map |
| 20 | 13.4 | 67.0 | 21 | AA268427 | AA268427 Human gen |
| 21 | 13.4 | 67.0 | 24 | AA111800 | AA111800 Diphtheri |
| 22 | 13.4 | 67.0 | 25 | AC149803 | AC149803 Human mic |
| 23 | 13.4 | 67.0 | 25 | AC181436 | AC181436 Human mic |

| | | | | | | |
|----|------|------|----|---|-----------|---------------------|
| 24 | 13.4 | 67.0 | 40 | 3 | AA289550 | AA289550 Human LAM |
| 25 | 13.4 | 67.0 | 50 | 6 | AB206362 | AB206362 Human leu |
| 26 | 13.4 | 67.0 | 50 | 6 | AB206165 | AB206165 Human leu |
| 27 | 13.2 | 66.0 | 20 | 3 | AAA27533 | AAA27533 Fas ligand |
| 28 | 13.2 | 66.0 | 22 | 6 | ABL44778 | ABL44778 Human chr |
| 29 | 13.2 | 66.0 | 22 | 6 | ABL44616 | ABL44616 Human chr |
| 30 | 13.2 | 66.0 | 25 | 8 | ACT15904 | ACT15904 Human mic |
| 31 | 13.2 | 66.0 | 30 | 2 | AA223278 | AA223278 Mouse str |
| 32 | 13.2 | 66.0 | 31 | 2 | AA223280 | AA223280 Mouse str |
| 33 | 13.2 | 66.0 | 50 | 4 | AA130677 | AA130677 Human SNP |
| 34 | 13.2 | 66.0 | 50 | 5 | AB100185 | AB100185 Human sll |
| 35 | 13 | 65.0 | 25 | 8 | ACT186815 | ACT186815 Human mic |
| 36 | 13 | 65.0 | 47 | 3 | AA266065 | AA266065 Human map |
| 37 | 12.8 | 64.0 | 21 | 2 | AA228619 | AA228619 Human par |
| 38 | 12.8 | 64.0 | 21 | 6 | AB259556 | AB259556 Real-time |
| 39 | 12.8 | 64.0 | 22 | 6 | AAQ46288 | AAQ46288 NF-1 gene |
| 40 | 12.8 | 64.0 | 24 | 2 | AA046948 | AA046948 Human neu |
| 41 | 12.8 | 64.0 | 24 | 4 | AA05047 | AA05047 Neurofibr |
| 42 | 12.8 | 64.0 | 25 | 8 | AC119405 | AC119405 Human mic |
| 43 | 12.8 | 64.0 | 25 | 8 | AC187928 | AC187928 Human mic |
| 44 | 12.8 | 64.0 | 25 | 8 | AC170105 | AC170105 Human mic |
| 45 | 12.8 | 64.0 | 25 | 8 | AC170105 | AC170105 Human mic |

ALIGNMENTS

RESULT 1
ADB25655
ID ADB25655 standard, DNA; 20 BP.

AC ADB25655;
DT 20-NOV-2003 (first entry)

XX Human connective tissue growth factor antisense oligo DNA (seqid 48).
XX antisense; human; ss; connective tissue growth factor; CTGF;
XX chromosome 6q23.1; cytofact; fibroblast inducible secreted protein;
XX fisp-12; NOV2;
XX insulin-like growth factor binding protein-related protein 2; IGFBP-IP2;
XX IGFBP-8; Hcs24; ecogenin; acute lymphoblastic leukemia; gene therapy;
XX hyperproliferative disorder; cancer; pulmonary fibrosis; renal fibrosis;
XX scleroderma; atherosclerosis; cytostatic; dermatological;
XX antiarteriosclerotic.

XX Homo sapiens.

XX Key Location/Qualifiers
XX modified_base 1..20
XX /tag= a
XX /mod_base= OTHER
XX /note= "OTHER= phosphorothioate backbone, where 1-5 and
16-20 are 2' methoxyethyl nucleotides. All cytidines are
5-methylcytidines"

XX NO200305340-A2.

XX 03-JUL-2003.

XX 09-DEC-2002; 2002MO-US038618.

XX 10-DEC-2001; 2001US-0006191.

XX (ISIS-) ISIS PHARM INC.

XX Garde WA, Walt AT;

XX WPI; 2003-5559091/52.

XX New antisense oligonucleotides for modulating connective tissue growth
factor expression, particularly useful for treating cancers (e.g. breast
or prostate cancer), pulmonary or renal fibrosis, scleroderma or

PT atherosclerosis.
 XX
 PS Claim 3; Page 85; 139pp; English.
 CC This invention relates to novel methods for modulating the expression of
 CC connective tissue growth factor (CTGF) by antisense oligonucleotides.
 CC CTGF has been mapped to human chromosome region 6q23.1, and is also known
 CC as ctgfact, fibroblast inducible secreted protein, fisp-12, NOV2,
 CC insulin-like growth factor binding protein-related protein 2, IGFBP-rp2,
 CC IGFBP-8, Hcs24 and ecogenin. It is known to stimulate DNA synthesis and
 CC promote chemotaxis of fibroblasts, however, it is also upregulated in
 CC acute lymphoblastic leukaemia and in tumour or endothelial cells
 CC associated with the vasculature. Accordingly, antisense oligonucleotides
 CC that inhibit the expression of CTGF in cells or tissues can be used in
 CC gene therapy to treat various conditions including hyperproliferative
 CC disorders (particularly cancer, e.g. breast, prostate or renal cancer),
 CC pulmonary fibrosis, renal fibrosis, scleroderma and atherosclerosis. As
 CC such, the present invention describes these antisense oligos as having
 CC cytostatic, dermatological and antiarteriosclerotic activities. This
 CC oligonucleotide sequence is a chimeric phosphorothioate antisense oligo
 CC with 2' MOE wings and a deoxy gap, which is used to inhibit expression of
 CC human CTGF of the invention.
 CC
 XX
 SQ Sequence 20 BP; 7 A; 3 C; 3 G; 7 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.1; Mismatches 0; Gaps 0;
 Matches 20; Conservative 0; Indels 0; Gaps 0;
 QY 1 AAACATGTAACCTTTGGTCA 20
 Db 1 AAACATGTAACCTTTGGTCA 20
 RESULT 2
 ADB25671
 ID ADB25671 standard; DNA; 20 BP.
 XX
 AC ADB25671;
 XX
 DT 20-NOV-2003 (first entry)
 DE Human connective tissue growth factor antisense oligo DNA (SeqID 64).
 XX
 KW antisense; human; ss; connective tissue growth factor; CTGF;
 KW chromosome 6q23.1; ctgfact; fibroblast inducible secreted protein;
 KW fisp-12; NOV2;
 KW insulin-like growth factor binding protein-related protein 2; IGFBP-rp2;
 KW IGFBP-8; Hcs24; ecogenin; acute lymphoblastic leukaemia; gene therapy;
 KW hyperproliferative disorder; cancer; pulmonary fibrosis; renal fibrosis;
 KW scleroderma; atherosclerosis; cytostatic; dermatological;
 KW antiarteriosclerotic.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT modified_base 1..20
 FT /*tag= a
 FT /mod_base= OTHER
 FT /note= "OTHER= phosphorothioate backbone, where 1-5 and
 FT 16-20 are 2' methoxyethyl nucleotides. All cytidines are
 FT 5-methylcytidines"
 XX
 PN WO2003053340-A2.
 XX
 PD 03-JUL-2003.
 XX
 PF 09-DEC-2002; 2002WC-US038618.
 XX
 PR 10-DEC-2001; 2001US-00006191.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX

PI Gaarde WA, Watt AT;
 XX
 DR WPI; 2003-559091/52.
 XX
 PT New antisense oligonucleotides for modulating connective tissue growth
 PT factor expression, particularly useful for treating cancers (e.g. breast
 PT or prostate cancer), pulmonary or renal fibrosis, scleroderma or
 PT atherosclerosis.
 XX
 PS Claim 3; Page 85; 139pp; English.
 CC This invention relates to novel methods for modulating the expression of
 CC connective tissue growth factor (CTGF) by antisense oligonucleotides.
 CC CTGF has been mapped to human chromosome region 6q23.1, and is also known
 CC as ctgfact, fibroblast inducible secreted protein, fisp-12, NOV2,
 CC insulin-like growth factor binding protein-related protein 2, IGFBP-rp2,
 CC IGFBP-8, Hcs24 and ecogenin. It is known to stimulate DNA synthesis and
 CC promote chemotaxis of fibroblasts, however, it is also upregulated in
 CC acute lymphoblastic leukaemia and in tumour or endothelial cells
 CC associated with the vasculature. Accordingly, antisense oligonucleotides
 CC that inhibit the expression of CTGF in cells or tissues can be used in
 CC gene therapy to treat various conditions including hyperproliferative
 CC disorders (particularly cancer, e.g. breast, prostate or renal cancer),
 CC pulmonary fibrosis, renal fibrosis, scleroderma and atherosclerosis. As
 CC such, the present invention describes these antisense oligos as having
 CC cytostatic, dermatological and antiarteriosclerotic activities. This
 CC oligonucleotide sequence is a chimeric phosphorothioate antisense oligo
 CC with 2' MOE wings and a deoxy gap, which is used to inhibit expression of
 CC human CTGF of the invention.
 CC
 XX
 SQ Sequence 20 BP; 6 A; 4 C; 3 G; 7 T; 0 U; 0 Other;
 Query Match 95.0%; Score 19; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 9.6;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAACATGTAACCTTTGGTGC 19
 Db 2 AAACATGTAACCTTTGGTGC 20
 RESULT 3
 ACT156439
 ID ACT156439 standard; DNA; 25 BP.
 XX
 AC ACT156439;
 XX
 DT 13-OCT-2003 (first entry)
 DE Human microarray DNA oligonucleotide SEQ ID NO 56430.
 XX
 KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
 KW genetic variation; biallelic marker; polymorphism; human;
 KW cross-species comparison.
 XX
 OS Homo sapiens.
 XX
 PN US2003104410-A1.
 XX
 PD 05-JUN-2003.
 XX
 PF 15-MAR-2002; 2002US-00098263.
 XX
 PR 16-MAR-2001; 2001US-0276759P.
 XX
 PA (AFFY-) AFFYMETRIX INC.
 XX
 PI Miltmann MP;
 XX
 DR WPI; 2003-567953/53.
 XX
 PT New array of nucleic acid probes, useful for in situ hybridization, in
 PT Southern, Northern or dot-blot hybridization to identify or detect the

PT sequence or specific mutations of any gene.
 XX Claim 1; SEQ ID NO 56430; 9pp; English.
 PS
 CC The invention discloses a microarray comprising a plurality of nucleic
 CC acid probes including one of 2,018,500 fully defined sequences, or its
 CC perfect match, perfect mismatch, antisense match or antisense mismatch.
 CC Also disclosed is a method of gene expression analysis. The array is used
 CC in monitoring gene expression levels by hybridisation to a DNA library,
 CC in analysis of genetic variation or in hybridisation of tag-labelled
 CC compounds. The nucleic acid probes are specifically designed for analysis
 CC of at least one target sequence. The method of analysis comprises
 CC hybridising at least one or more nucleic acids to at least two or more
 CC nucleic acid probes and detecting the hybridisation. The nucleic acid
 CC probes are attached to a solid support. The analysis comprises monitoring
 CC gene expression levels, identifying allelic markers or polymorphisms,
 CC or family members of a gene and a cross-species comparison. Each of the
 CC nucleic acids further comprises a tag sequence. The array of nucleic acid
 CC probes is useful in situ hybridisation, in Southern, Northern or dot-
 CC blot hybridisation to identify or detect the sequence or specific
 CC mutations of any gene, in mapping the 5' terminus of mRNA molecules by
 CC primer extensions or in screening cDNA or genomic libraries or subclones
 CC for additional subclones containing segments of DNA that have been
 CC isolated and previously sequenced. The sequence presented is one of the
 CC nucleic acid probes incorporated in the microarray. Note: The sequence
 CC data for this patent can also be obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html
 CC
 XX
 SQ Sequence 25 BP; 6 A; 4 C; 4 G; 11 T; 0 U; 0 Other;
 Query Match 76.0%; Score 15.2; DB 8; Length 25;
 Best Local Similarity 85.0%; Pred. No. 7.5e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 AACATCTACTTTTGTCGA 20
 Db 1 AGACATCTACTTTTTCGA 20
 ID ADB25654 standard; DNA; 20 BP.
 XX
 AC ADB25654;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human connective tissue growth factor antisense oligo DNA (SeqID 47).
 XX
 KW antisense; human; ss; connective tissue growth factor; CTGF;
 KW chromosome 6q23.1; ctgfact; fibroblast inducible secreted protein;
 KW fisp-12; NOV2;
 KW insulin-like growth factor binding protein-related protein 2; IGFBP-
 KW IGFBP-8; Hcs24; ecogenin; acute lymphoblastic leukaemia; gene therapy;
 KW hyperproliferative disorder; cancer; pulmonary fibrosis; renal fibrosis;
 KW scleroderma; atherosclerosis; cytosratic; dermatological;
 KW antiarteriosclerotic.
 KW
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT modified_base 1..20
 FT /*tag= a
 FT /mod_base= OTHER
 FT /note= "OTHER= phosphorothioate backbone, where 1-5 and
 FT 16-20 are 2' methoxyethyl nucleotides. All cytidines are
 FT 5-methylcytidines"
 XX
 XX WO2003053340-A2.
 XX 03-JUN-2003.
 XX 09-DEC-2002; 2002WO-US038618.

XX
 PR 10-DEC-2001; 2001US-00006191.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Gaarde WA, Watt AT;
 XX
 DR WPI; 2003-559091/52.
 XX
 PI New antisense oligonucleotides for modulating connective tissue growth
 PI factor expression, particularly useful for treating cancers (e.g. breast
 PI or prostate cancer), pulmonary or renal fibrosis, scleroderma or
 PI atherosclerosis.
 PS
 PS Claim 3; Page 85; 139pp; English.
 XX
 CC This invention relates to novel methods for modulating the expression of
 CC connective tissue growth factor (CTGF) by antisense oligonucleotides.
 CC CTGF has been mapped to human chromosome region 6q23.1, and is also known
 CC as ctgfact, fibroblast inducible secreted protein, fisp-12, NOV2,
 CC insulin-like growth factor binding protein-related protein 2, IGFBP-
 CC IGFBP-8, Hcs24 and ecogenin. It is known to stimulate DNA synthesis and
 CC promote chemotaxis of fibroblasts, however, it is also upregulated in
 CC acute lymphoblastic leukaemia and in tumour or endothelial cells
 CC associated with the vasculature. Accordingly, antisense oligonucleotides
 CC that inhibit the expression of CTGF in cells or tissues can be used in
 CC gene therapy to treat various conditions including hyperproliferative
 CC disorders (particularly cancer, e.g. breast, prostate or renal cancer),
 CC pulmonary fibrosis, renal fibrosis, scleroderma and atherosclerosis. As
 CC such, the present invention describes these antisense oligos as having
 CC cytostatic, dermatological and antiarteriosclerotic activities. This
 CC oligonucleotide sequence is a chimeric phosphorothioate antisense oligo
 CC with 2' MOE wings and a deoxy gap, which is used to inhibit expression of
 CC human CTGF of the invention.
 CC
 XX
 SQ Sequence 20 BP; 4 A; 5 C; 3 G; 8 T; 0 U; 0 Other;
 Query Match 75.0%; Score 15; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 9.3e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 TGTACTTTTGTCGA 20
 Db 1 TGTACTTTTGTCGA 15
 ID ACI49802 standard; DNA; 25 BP.
 XX
 AC ACI49802;
 XX
 DT 13-OCT-2003 (first entry)
 XX
 DE Human microarray DNA oligonucleotide SEQ ID NO 49793.
 XX
 KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
 KW genetic variation; allelic marker; polymorphism; human;
 KW cross-species comparison.
 KW
 OS Homo sapiens.
 XX
 OS US2003104410-A1.
 XX
 PD 05-JUN-2003.
 XX
 PF 15-MAR-2002; 2002US-00098263.
 XX
 PR 16-MAR-2001; 2001US-0276759P.
 XX
 PA (AFFY-) AFFYMETRIX INC.
 XX
 PI Miltmann MP;

XX WPI; 2003-567953/53.
 XX New array of nucleic acid probes, useful for in situ hybridization, in
 PT Southern, Northern or dot-blot hybridization to identify or detect the
 PT sequence or specific mutations of any gene.
 XX
 PS Claim 1; SEQ ID NO 49793; 9pp; English.

XX The invention discloses a microarray comprising a plurality of nucleic
 CC acid probes including one of 2,018,500 fully defined sequences, or its
 CC perfect match, perfect mismatch, antisense match or antisense mismatch.
 CC Also disclosed is a method of gene expression analysis. The array is used
 CC in monitoring gene expression levels by hybridization to a DNA library,
 CC in analysis of genetic variation or in hybridization of tag-labelled
 CC compounds. The nucleic acid probes are specifically designed for analysis
 CC of at least one target sequence. The method of analysis comprises
 CC hybridizing at least one or more nucleic acids to at least two or more
 CC nucleic acid probes and detecting the hybridization. The nucleic acid
 CC probes are attached to a solid support. The analysis comprises monitoring
 CC gene expression levels, identifying allelic markers or polymorphisms,
 CC or family members of a gene and a cross-species comparison. Each of the
 CC nucleic acids further comprises a tag sequence. The array of nucleic acid
 CC probes is useful in situ hybridization, in Southern, Northern or dot-
 CC blot hybridization to identify or detect the sequence or specific
 CC mutations of any gene, in mapping the 5' terminus of mRNA molecules by
 CC primer extensions or in screening cDNA or genomic libraries or subclones
 CC for additional subclones containing segments of DNA that have been
 CC isolated and previously sequenced. The sequence presented is one of the
 CC nucleic acid probes incorporated in the microarray. Note: The sequence
 CC data for this patent can also be obtained in electronic format directly
 CC from USPRO at seqdata.uspto.gov/sequence.html

XX Sequence 25 BP; 8 A; 4 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 75.0%; Score 15; DB 8; Length 25;
 Best Local Similarity 100.0%; Pred. No. 9.5e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ACATGTAACCTTTGG 17
 |||||
 DB 15 ACATGTAACCTTTGG 1

RESULT 6
 ABQ93618/c
 ID ABQ93618 standard; DNA; 25 BP.

XX AEO93618;

DT 16-OCT-2002 (first entry)

XX Human DISC1/DISC2 PCR primer disc21 r2.

XX Human; disrupted in Schizophrenia 1; DISC1; neuroleptic; gene therapy;
 KW neuropsychiatric disorder; schizoaffective disorder; bipolar disorder;
 KM unipolar affective disorder; adolescent conduct disorder; schizophrenia;
 KM PCR; primer; ss.

XX Homo sapiens.

XX W0200258637-A2.

XX 01-AUG-2002.

XX 23-JAN-2002; 2002WO-US002186.

XX 24-JAN-2001; 2001US-00770107.

XX (MILL-) MILLENIUM PHARM INC.

PI Meyer JM, Barrington-Martin R, Parker A, Barnes GT;

DR WPI; 2002-590791/63.
 XX New human Disrupted-in-Schizophrenia (DISC) 1 and DISC2 genes containing
 PT single nucleotide polymorphisms, useful for preventing or treating
 PT neuropsychiatric disorders e.g. schizophrenia.
 XX
 PS Claim 17; Fig 4; 169pp; English.

XX The invention relates to a novel Disrupted-in-Schizophrenia (DISC) 1
 CC allelic variant polymorphisms. The polymorphisms of the invention have
 CC neuroleptic activity. The polymorphisms may have a use in gene therapy.
 CC DISC1 or DISC2 nucleic acid molecules are useful for diagnosing or
 CC treating a subject having a disease or disorder associated with specific
 CC DISC1 or DISC2 alleles and/or aberrant DISC1 expression or activity e.g.
 CC neuropsychiatric disorder such as schizoaffective, bipolar, unipolar
 CC affective or adolescent conduct disorder or schizophrenia. Similarly, the
 CC compound that inhibits DISC1 protein activity may be used in the method
 CC for treating such neuropsychiatric disorders. The sequences shown in
 CC ABQ93575-ABQ93658 represent the PCR primers used in the invention to
 CC amplify the sequences of DISC2 and DISC2

XX Sequence 25 BP; 5 A; 6 C; 6 G; 8 T; 0 U; 0 Other;

Query Match 74.0%; Score 14.8; DB 6; Length 25;
 Best Local Similarity 88.9%; Pred. No. 1.2e+03;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AACATGTAACCTTTGGT 18
 |||||
 DB 21 AACATGTAACCTTTGGT 4

RESULT 7
 ACT15905
 ID ACT15905 standard; DNA; 25 BP.

XX ACT15905;

DT 13-OCT-2003 (first entry)

XX Human microarray DNA oligonucleotide SEQ ID NO 15896.

XX EST; ss; probe; expressed sequence tag; microarray; gene expression;
 KW genetic variation; allelic marker; polymorphism; human;
 KM cross-species comparison.

XX Homo sapiens.

XX US2003104410-A1.

XX 05-JUN-2003.

XX 15-MAR-2002; 2002US-00098263.

XX 16-MAR-2001; 2001US-0276759P.

XX (AFFY-) AFFYMETRIX INC.

XX Miltmann MP;

XX WPI; 2003-567953/53.

XX New array of nucleic acid probes, useful for in situ hybridization, in
 PT Southern, Northern or dot-blot hybridization to identify or detect the
 PT sequence or specific mutations of any gene.

XX Claim 1; SEQ ID NO 15896; 9pp; English.

XX The invention discloses a microarray comprising a plurality of nucleic
 CC acid probes including one of 2,018,500 fully defined sequences, or its
 CC perfect match, perfect mismatch, antisense match or antisense mismatch.
 CC Also disclosed is a method of gene expression analysis. The array is used
 CC in monitoring gene expression levels by hybridization to a DNA library,

CC in analysis of genetic variation or in hybridisation of tag-labelled
 CC compounds. The nucleic acid probes are specifically designed for analysis
 CC of at least one target sequence. The method of analysis comprises
 CC hybridising at least one or more nucleic acids to at least two or more
 CC nucleic acid probes and detecting the hybridisation. The nucleic acid
 CC probes are attached to a solid support. The analysis comprises monitoring
 CC gene expression levels, identifying diallelic markers or polymorphisms,
 CC or family members of a gene and a cross-species comparison. Each of the
 CC nucleic acids further comprises a tag sequence. The array of nucleic acid
 CC probes is useful in situ hybridisation, in Southern, Northern or dot-
 CC blot hybridisation to identify or detect the sequence or specific
 CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
 CC primer extensions or in screening cDNA or genomic libraries or subclones
 CC for additional subclones containing segments of DNA that have been
 CC isolated and previously sequenced. The sequence presented is one of the
 CC nucleic acid probes incorporated in the microarray. Note: The sequence
 CC data for this patent can also be obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html

SO Sequence 25 BP; 5 A; 5 C; 4 G; 11 T; 0 U; 0 Other;

Query Match 74.0%; Score 14.8; DB 8; Length 25;
 Best Local Similarity 88.9%; Pred. No. 1.2e+03;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ACATGTAACCTTTGGTCA 20
 |||||
 Db 4 ACATGTAACCTTTGGTCA 21

RESULT 8
 ACK05946
 ID ACK05946 standard; DNA; 25 BP.

AC ACK05946;

DT 14-OCT-2003 (first entry)

DE Human microarray DNA oligonucleotide SEQ ID NO 105927.

KW EST; ss; probe; expressed sequence tag; microarray; gene expression;

KM genetic variation; diallelic marker; polymorphism; human;

KW cross-species comparison.

OS Homo sapiens.

PN US2003104410-A1.

PD 05-JUN-2003.

PF 15-MAR-2002; 2002US-00098263.

PR 16-MAR-2001; 2001US-0276759P.

PA (AFFY-) AFFYMETRIX INC.

PI Miltmann MP;

PT WPI; 2003-567953/53.

PT New array of nucleic acid probes, useful for in situ hybridization, in
 PT Southern, Northern or dot-blot hybridization to identify or detect the
 PT sequence or specific mutations of any gene.

PS Claim 1; SEQ ID NO 105927; 9pp; English.

CC The invention discloses a microarray comprising a plurality of nucleic
 CC acid probes including one of 2,018,500 fully defined sequences, or its
 CC perfect match, perfect mismatch, antisense match or antisense mismatch.
 CC Also disclosed is a method of gene expression analysis. The array is used
 CC in monitoring gene expression levels by hybridisation to a DNA library,
 CC in analysis of genetic variation or in hybridisation of tag-labelled
 CC compounds. The nucleic acid probes are specifically designed for analysis

CC of at least one target sequence. The method of analysis comprises
 CC hybridising at least one or more nucleic acids to at least two or more
 CC nucleic acid probes and detecting the hybridisation. The nucleic acid
 CC probes are attached to a solid support. The analysis comprises monitoring
 CC gene expression levels, identifying diallelic markers or polymorphisms,
 CC or family members of a gene and a cross-species comparison. Each of the
 CC nucleic acids further comprises a tag sequence. The array of nucleic acid
 CC probes is useful in situ hybridisation, in Southern, Northern or dot-
 CC blot hybridisation to identify or detect the sequence or specific
 CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
 CC primer extensions or in screening cDNA or genomic libraries or subclones
 CC for additional subclones containing segments of DNA that have been
 CC isolated and previously sequenced. The sequence presented is one of the
 CC nucleic acid probes incorporated in the microarray. Note: The sequence
 CC data for this patent can also be obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html

SO Sequence 25 BP; 9 A; 4 C; 5 G; 7 T; 0 U; 0 Other;

Query Match 71.0%; Score 14.2; DB 8; Length 25;
 Best Local Similarity 84.2%; Pred. No. 2.4e+03;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 AACGTGTAACCTTTGGTCA 20
 |||||
 Db 6 AAAAGTAACCTTTGGTCA 24

RESULT 9
 ADB25670
 ID ADB25670 standard; DNA; 20 BP.

AC ADB25670;

DT 20-NOV-2003 (first entry)

DE Human connective tissue growth factor antisense oligo DNA (SeqID 63).

KW antisense; human; ss; connective tissue growth factor; CTGF;

KM chromosome 6q23.1; ctgofact; fibroblast inducible secreted protein;

KW fisp-12; NOV2;

KW insulin-like growth factor binding protein-related protein 2; IGFBP-1P2;

KW IGFBP-8; Hcs24; ecogenin; acute lymphoblastic leukaemia; gene therapy;

KW hyperproliferative disorder; cancer; pulmonary fibrosis; renal fibrosis;

KW scleroderma; atherosclerosis; cyrostatic; dermatological;

KW antiarteriosclerotic.

OS Homo sapiens.

PN WO2003053340-A2.

PD 03-JUL-2003.

PF 09-DEC-2002; 2002WO-US038618.

PR 10-DEC-2001; 2001US-00006191.

PA (ISIS-) ISIS PHARM INC.

PI Gaarde WA, Walt AT;

PT WPI; 2003-559091/52.

CC New antisense oligonucleotides for modulating connective tissue growth
 CC factor expression, particularly useful for treating cancers (e.g. breast

PT or prostate cancer), pulmonary or renal fibrosis, scleroderma or
 PT atherosclerosis.
 XX
 PS Claim 3, Page 85, 139pp; English.
 XX
 CC This invention relates to novel methods for modulating the expression of
 CC connective tissue growth factor (CTGF) by antisense oligonucleotides.
 CC CTGF has been mapped to human chromosome region 6q23.1 and is also known
 CC as ccgrefact, fibroblast inducible secreted protein, fibp-12, NOV2,
 CC insulin-like growth factor binding protein-related protein 2, IGFBP-
 CC IGFBP-8, Hs824 and ecogenin. It is known to stimulate DNA synthesis and
 CC promote chemotaxis of fibroblasts, however, it is also upregulated in
 CC acute lymphoblastic leukaemia and in tumour or endothelial cells
 CC associated with the vasculature. Accordingly, antisense oligonucleotides
 CC that inhibit the expression of CTGF in cells or tissues can be used in
 CC gene therapy to treat various conditions including hyperproliferative
 CC disorders (particularly cancer, e.g. breast, prostate or renal cancer),
 CC pulmonary fibrosis, renal fibrosis, scleroderma and atherosclerosis. As
 CC such, the present invention describes these antisense oligos as having
 CC cytostatic, dermatological and antiarteriosclerotic activities. This
 CC oligonucleotide sequence is a chimeric phosphorothioate antisense oligo
 CC with 2' MOE wings and a deoxy gap, which is used to inhibit expression of
 CC human CTGF of the invention.
 CC
 SO Sequence 20 BP; 4 A; 5 C; 3 G; 6 T; 0 U; 0 Other;
 Query Match 70.0%; Score 14; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.9e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 GTAACTTTGGTCA 20
 DB 1 GTAACTTTGGTCA 14
 ACII7811/C
 ID ACII7811 standard; DNA; 25 BP.
 ACII7811;
 ACII7811;
 13-OCT-2003 (first entry)
 DE Human microarray DNA oligonucleotide SEQ ID NO 57645.
 KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
 KW genetic variation; diallelic marker; polymorphism; human;
 KW cross-species comparison.
 OS Homo sapiens.
 XX
 PN US2003104410-A1.
 PD 05-JUN-2003.
 PF 15-MAR-2002; 2002US-00098263.
 PR 16-MAR-2001; 2001US-0276759P.
 PA (AFY-) AFFYMETRIX INC.
 PI Mittmann MP;
 DR WPI; 2003-567953/53.
 XX
 PT New array of nucleic acid probes, useful for in situ hybridization, in
 PT Southern, Northern or dot-blot hybridization to identify or detect the
 PT sequence or specific mutations of any gene.
 XX
 PS Claim 1; SEQ ID NO 57645; 9pp; English.
 CC The invention discloses a microarray comprising a plurality of nucleic
 CC acid probes including one of 2,018,500 fully defined sequences, or its

CC perfect match, perfect mismatch, antisense match or antisense mismatch.
 CC Also disclosed is a method of gene expression analysis. The array is used
 CC in monitoring gene expression levels by hybridization to a DNA library,
 CC in analysis of genetic variation or in hybridization of tag-labelled
 CC compounds. The nucleic acid probes are specifically designed for analysis
 CC of at least one target sequence. The method of analysis comprises
 CC hybridizing at least one or more nucleic acids to at least two or more
 CC nucleic acid probes and detecting the hybridization. The nucleic acid
 CC probes are attached to a solid support. The analysis comprises monitoring
 CC gene expression levels, identifying diallelic markers or polymorphisms,
 CC or family members of a gene and a cross-species comparison. Each of the
 CC nucleic acids further comprises a tag sequence. The array of nucleic acid
 CC probes is useful in in situ hybridization, in Southern, Northern or dot-
 CC blot hybridization to identify or detect the sequence or specific
 CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
 CC primer extensions or in screening cDNA or genomic libraries or subclones
 CC for additional subclones containing segments of DNA that have been
 CC isolated and previously sequenced. The sequence presented is one of the
 CC nucleic acid probes incorporated in the microarray. Note: The sequence
 CC data for this patent can also be obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html
 CC
 SO Sequence 25 BP; 8 A; 4 C; 6 G; 7 T; 0 U; 0 Other;
 Query Match 68.0%; Score 13.6; DB 8; Length 25;
 Best Local Similarity 80.0%; Pred. No. 4.7e+03;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AAACATGTAACCTTTCTTA 20
 DB 23 AAACATGTAACCTTTCTTA 4
 ACII7811/C
 ID ACII7811 standard; DNA; 25 BP.
 ACII7811;
 ACII7811;
 13-OCT-2003 (first entry)
 DE Human microarray DNA oligonucleotide SEQ ID NO 17802.
 KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
 KW genetic variation; diallelic marker; polymorphism; human;
 KW cross-species comparison.
 OS Homo sapiens.
 XX
 PN US2003104410-A1.
 PD 05-JUN-2003.
 PF 15-MAR-2002; 2002US-00098263.
 PR 16-MAR-2001; 2001US-0276759P.
 PA (AFY-) AFFYMETRIX INC.
 PI Mittmann MP;
 DR WPI; 2003-567953/53.
 XX
 PT New array of nucleic acid probes, useful for in situ hybridization, in
 PT Southern, Northern or dot-blot hybridization to identify or detect the
 PT sequence or specific mutations of any gene.
 XX
 PS Claim 1; SEQ ID NO 17802; 9pp; English.
 CC The invention discloses a microarray comprising a plurality of nucleic
 CC acid probes including one of 2,018,500 fully defined sequences, or its
 CC perfect match, perfect mismatch, antisense match or antisense mismatch.
 CC Also disclosed is a method of gene expression analysis. The array is used

in monitoring gene expression levels by hybridisation to a DNA library, in analysis of genetic variation or in hybridisation of tag-labelled compounds. The nucleic acid probes are specifically designed for analysis of at least one target sequence. The method of analysis comprises hybridising at least one or more nucleic acids to at least two or more nucleic acid probes and detecting the hybridisation. The nucleic acid probes are attached to a solid support. The analysis comprises monitoring gene expression levels, identifying diallelic markers or polymorphisms, or family members of a gene and a cross-species comparison. Each of the nucleic acids further comprises a tag sequence. The array of nucleic acid probes is useful in situ hybridisation, in Southern, Northern or dot-blot hybridisation to identify or detect the sequence or specific mutations of any gene, in mapping the 5' termini of mRNA molecules by primer extensions or in screening cDNA or genomic libraries or subclones for additional subclones containing segments of DNA that have been isolated and previously sequenced. The sequence presented is one of the nucleic acid probes incorporated in the microarray. Note: The sequence data for this patent can also be obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html

Sequence 25 BP; 7 A; 4 C; 7 G; 7 T; 0 U; 0 Other;

Query Match 68.0%; Score 13.6; DB 8; Length 25;
Best Local Similarity 80.0%; Pred. No. 4.7e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAACATGTAAGTCTTGGTCA 20
DB 20 AAACATGTAAGTCTTGGTCA 1

RESULT 12

AC156438 standard; DNA; 25 BP.

AC156438;

13-OCT-2003 (first entry)

Human microarray DNA oligonucleotide SEQ ID NO 56429.

EST, ss: probe; expressed sequence tag; microarray; gene expression; genetic variation; diallelic marker; polymorphism; human; cross-species comparison.

Homo sapiens.

US2003104410-A1.

05-JUN-2003.

15-MAR-2002; 2002US-00098263.

16-MAR-2001; 2001US-0276759P.

(AFYV-) AFFYMETRIX INC.

Mittmann MF;

WPI; 2003-567953/53.

New array of nucleic acid probes, useful for in situ hybridization, in Southern, Northern or dot-blot hybridization to identify or detect the sequence or specific mutations of any gene.

Claim 1; SEQ ID NO 56429; 9pp; English.

The invention discloses a microarray comprising a plurality of nucleic acid probes including one of 2,018,500 fully defined sequences, or its perfect match, perfect mismatch, antisense match or antisense mismatch. Also disclosed is a method of gene expression analysis. The array is used in monitoring gene expression levels by hybridisation to a DNA library, in analysis of genetic variation or in hybridisation of tag-labelled

compounds. The nucleic acid probes are specifically designed for analysis of at least one target sequence. The method of analysis comprises hybridising at least one or more nucleic acids to at least two or more nucleic acid probes and detecting the hybridisation. The nucleic acid probes are attached to a solid support. The analysis comprises monitoring gene expression levels, identifying diallelic markers or polymorphisms, or family members of a gene and a cross-species comparison. Each of the nucleic acids further comprises a tag sequence. The array of nucleic acid probes is useful in situ hybridisation, in Southern, Northern or dot-blot hybridisation to identify or detect the sequence or specific mutations of any gene, in mapping the 5' termini of mRNA molecules by primer extensions or in screening cDNA or genomic libraries or subclones for additional subclones containing segments of DNA that have been isolated and previously sequenced. The sequence presented is one of the nucleic acid probes incorporated in the microarray. Note: The sequence data for this patent can also be obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html

Sequence 25 BP; 7 A; 4 C; 4 G; 10 T; 0 U; 0 Other;

Query Match 68.0%; Score 13.6; DB 8; Length 25;
Best Local Similarity 80.0%; Pred. No. 4.7e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAACATGTAAGTCTTGGTCA 20
DB 1 AGACATGTAAGTCTTGGTCA 20

RESULT 13

ADD20567/C standard; DNA; 25 BP.

ADD20567;

15-JAN-2004 (first entry)

Oreochromis niloticus microsatellite primer SEQ ID NO:1202.

single nucleotide polymorphism; SNP; fish; Salmo salar; Oreochromis niloticus; Atlantic halibut; microsatellite; cod; polymorphic site; seabass; salmonidae; tilapia; rainbow trout; halibut; detection; primer; ss.

Synthetic.

Oreochromis niloticus.

WO2003060160-A2.

24-JUL-2003.

17-JAN-2003; 2003WO-1B000112.

18-JAN-2002; 2002US-0349950P.

16-AUG-2002; 2002US-0404200P.

(GENO-) GENOMAR ASA.

Lie O, Slettan A, Hoyum M, Lingaas F;

WPI; 2003-627388/59.

Novel isolated nucleic acid molecule comprising single nucleotide polymorphism associated with fish, useful for forming PCR primers which are used for detecting single nucleotide polymorphisms in fish nucleic acids.

Claim 18; SEQ ID NO 1202; 23pp; English.

The present invention describes an isolated nucleic acid (I) comprising a single nucleotide polymorphism (SNP) chosen from: (i) a nucleic acid of Salmo salar SNPs, Oreochromis niloticus SNPs or Atlantic halibut SNPs; and (ii) a nucleic acid having nucleotide sequence that hybridises to

(1), or its complement under highly stringent hybridisation conditions. Also described: (1) an isolated oligonucleotide (II) comprising at least 1 contiguous nucleotides of a nucleotide sequence of *S. salar* SNPs, *O. niloticus* SNPs, *O. niloticus* microsatellites, Atlantic halibut SNPs, cod polymorphic sites and seabass polymorphic sites, or their complement; (2) a primer pair (III) suitable for use in PCR, comprising two (II) capable of amplifying a nucleotide sequence chosen from *S. salar* SNPs and, *O. niloticus* SNPs, *O. niloticus* microsatellites, Atlantic halibut SNPs, cod polymorphic sites and seabass polymorphic sites; and determining (M1) the origin of fish sample comprising providing a parent genotype database comprising a collection of candidate parent genotypes, where each of the candidate parent genotype represents a distinct origin, and comparing a sample genotype to the parent genotype database, where a match between the sample genotype and one of the candidate parent genotype identifies to the origin of the sample. (M1) is useful for determining the origin of a fish sample such as family salmonidae, *S. salar*, tilapia, *O. niloticus*, rainbow trout, halibut, seabass and Atlantic cod. (II) is useful for detecting nucleic acid molecule comprising SNP in a sample, which involves contacting the sample containing nucleic acids with one or more SNPs, and identifying nucleic acid that hybridises to (II). (II) is useful for detecting nucleic acid molecule comprising a polymorphic sequence in a sample, comprising contacting the sample containing nucleic acids with one or more (II) which is derived from *O. niloticus* microsatellite, *O. niloticus* SNPs, Atlantic halibut SNPs, cod polymorphic sites or seabass polymorphic sites, and identifying a nucleic acid that hybridises to (II). (II) is useful for detecting nucleic acid molecule comprising a microsatellite sequence in sample. The present sequence is used in the exemplification of the present invention.

Sequence 25 BP; 9 A; 7 C; 3 G; 6 T; 0 U; 0 Other;

Query Match 68.0%; Score 13.6; DB 9; Length 25;
Best Local Similarity 80.0%; Pred. No. 4.7e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 AACATGTAAGTTTGTCGA 20
23 AGAGTGTAACTTGTCTCA 4

RESULT 14
AA068537

ID AA068537 standard; CDNA; 26 BP.

AC AA068537;

XX 25-MAR-2003 (revised)

DT 13-FEB-1995 (first entry)

XX B.thuringiensis 33kD delta-endotoxin N-terminal probe.

XX Insecticidal protein; delta-endotoxin; crystal; Coleoptera; Lepidoptera;

KM *Bacillus thuringiensis*, ss.

XX Synthetic.

XX WO9413785-A2.

XX 23-JUN-1994.

XX 13-DEC-1993; 93WO-US012144.

XX 15-DEC-1992; 92US-00991073.

XX (NOVO) NOVO-NORDISK ENTOTEC INC.

XX Liu C, Adams LF, Lufburrow PA, Thomas MD;

XX WPI; 1994-217865/26.

XX New *Bacillus thuringiensis* strains - which produce new delta-endotoxin
PT cpds used for the control of Lepidopteran and Coleopteran insect pests.

XX Example 9; Page 26; 47P; English.

XX The N-terminal sequence of a 33kD delta-endotoxin isolated from
CC B.thuringiensis EMCC0075 (NRRL B-21019) or EMCC0076 (NRRL B-21020) was
CC determined (see AAS59764). Based on this sequence a 26mer oligonucleotide
CC was designed for use as a probe (AA068537) for cloning the delta-
CC endotoxin gene. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 26 BP; 13 A; 2 C; 2 G; 9 T; 0 U; 0 Other;

Query Match 68.0%; Score 13.6; DB 2; Length 26;
Best Local Similarity 80.0%; Pred. No. 4.7e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 AACATGTAAGTTTGTCGA 20
4 AAACATATAAATTTTGATCA 23

RESULT 15
AAT51400

ID AAT51400 standard; DNA; 33 BP.

AC AAT51400;

XX 23-APR-1997 (first entry)

XX D. immitis VAS-like gene clone ndiVA726 sense primer BVVA sen.

XX Venom allergen antigen 5-like gene; VAS; helminth; parasite; ndiVA726;

KM vaccine; heartworm; Dirofilaria immitis; primer; PCR;

XX polymerase chain reaction; ss.

XX Synthetic.

XX WO9637218-A1.

XX 28-NOV-1996.

XX 23-MAY-1996; 96WO-US007709.

XX 23-MAY-1995; 95US-00450944.

XX (HESK-) HESKA CORP.

XX Tripp CA, Wisniewski N;

XX Nucleic acid encoding helminth venom allergen antigen 5-like protein -

PT pref. from Dirofilaria immitis or Onchocerca volvulus, useful esp. in

XX vaccines to prevent helminth infection.

XX Example 4; Page 68; 124P; English.

CC Sense primer BVVA sen (AAT51400) contg. a BamHI site, and antisense
CC primer BVVA ant (AAT51401) contg. a BclRI site, were used in the PCR
CC amplification of a 726-nucleotide DNA fragment, designated ndiVA726. This
CC DNA fragment corresponds to nucleotides 54-779 of the Dirofilaria immitis
CC nucleic acid molecule ndiVA833 (see also AAT51376), which encodes a venom
CC allergen antigen 5 (VAS)-like protein (AAM1479). The PCR product was
CC incorporated into a baculovirus vector to enable prodn. of recombinant
CC VAS-like protein in Spodoptera frugiperda Sf9 cells

Sequence 33 BP; 9 A; 5 C; 6 G; 13 T; 0 U; 0 Other;

Query Match 68.0%; Score 13.6; DB 2; Length 33;
Best Local Similarity 80.0%; Pred. No. 4.8e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 AACATGTAAGTTTGTCGA 20
|||||

Mon Mar 29 09:29:16 2004

us-10-006-191-48.1150.rng

Page 9

DB 13 AAATATGATACCTTTGGTTA 32

Search completed: March 27, 2004, 09:34:17
Job time : 341 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2004, 09:22:36 ; Search time 2406 Seconds
(without alignments)
248.231 Million cell updates/sec

Title: US-10-006-191-48

Sequence: 1 aaacatgtaactttgtgtca 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 138346

Minimum DB seg length: 0

Maximum DB seg length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estp1:*
7: em_estro:*
8: em_estc:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estum:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrt1:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 14.4 | 72.0 | 47 | 28 | B2354183 |
| 2 | 13.6 | 68.0 | 40 | 28 | BH853744 |
| 3 | 13.6 | 68.0 | 42 | 29 | AL942024 |
| 4 | 13.2 | 66.0 | 35 | 12 | B0044144 |

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 5 | 12.8 | 64.0 | 40 | 28 | A2796505 |
| 6 | 12.8 | 64.0 | 42 | 28 | BH911929 |
| 7 | 12.8 | 64.0 | 46 | 13 | B0595103 |
| 8 | 12.4 | 62.0 | 48 | 28 | CC456918 |
| 9 | 12.4 | 62.0 | 49 | 10 | AW395746 |
| 10 | 12.2 | 61.0 | 23 | 29 | TA264E040 |
| 11 | 12.2 | 61.0 | 40 | 29 | CC886423 |
| 12 | 12.2 | 61.0 | 47 | 29 | BX659565 |
| 13 | 12.2 | 61.0 | 48 | 28 | A2602539 |
| 14 | 12.2 | 61.0 | 49 | 28 | B2382085 |
| 15 | 12.2 | 61.0 | 49 | 29 | CC886429 |
| 16 | 12.2 | 61.0 | 30 | 28 | A2619314 |
| 17 | 12.2 | 60.0 | 32 | 28 | A2773045 |
| 18 | 12.2 | 60.0 | 32 | 28 | BH901742 |
| 19 | 12.2 | 60.0 | 32 | 28 | BH904527 |
| 20 | 12.2 | 60.0 | 34 | 28 | A2448375 |
| 21 | 12.2 | 60.0 | 41 | 14 | CF295817 |
| 22 | 12.2 | 60.0 | 42 | 29 | CC886626 |
| 23 | 12.2 | 60.0 | 46 | 28 | A2470558 |
| 24 | 12.2 | 60.0 | 46 | 28 | A2499560 |
| 25 | 12.2 | 60.0 | 46 | 28 | BH863928 |
| 26 | 12.2 | 60.0 | 49 | 14 | CF099660 |
| 27 | 11.8 | 59.0 | 34 | 9 | A2670672 |
| 28 | 11.8 | 59.0 | 36 | 9 | AV837544 |
| 29 | 11.8 | 59.0 | 37 | 28 | A2660403 |
| 30 | 11.8 | 59.0 | 38 | 28 | A2393785 |
| 31 | 11.8 | 59.0 | 40 | 28 | B2354783 |
| 32 | 11.8 | 59.0 | 40 | 28 | B2354833 |
| 33 | 11.8 | 59.0 | 41 | 28 | BH853687 |
| 34 | 11.8 | 59.0 | 43 | 29 | AL762323 |
| 35 | 11.8 | 59.0 | 46 | 9 | AA073235 |
| 36 | 11.6 | 58.0 | 28 | 28 | BH861790 |
| 37 | 11.6 | 58.0 | 29 | 28 | A2623171 |
| 38 | 11.6 | 58.0 | 37 | 28 | BH911930 |
| 39 | 11.6 | 58.0 | 37 | 28 | A2307043 |
| 40 | 11.6 | 58.0 | 39 | 28 | BH861726 |
| 41 | 11.6 | 58.0 | 39 | 28 | BH861727 |
| 42 | 11.6 | 58.0 | 40 | 29 | AL764914 |
| 43 | 11.6 | 58.0 | 41 | 28 | BH861762 |
| 44 | 11.6 | 58.0 | 42 | 13 | C01292 |
| 45 | 11.6 | 58.0 | 42 | 29 | CC887982 |

ALIGNMENTS

RESULT 1
B2354183/c
LOCUS
DEFINITION
B2354183
SALK_123321.19.20.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_123321.19.20.x, genomic
survey sequence.
ACCESSION
B2354183
VERSION
B2354183.1
GI:24945045
KEYWORDS
GSS.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Nucleotide
Arabidopsis thaliana
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 47)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,T.,
Shim,P., Zimmerman,J., and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379

Email: ecker@salik.edu
This is single pass sequence recovered from the left border of TDNA.

FEATURES

source

Class: TDNA tagged.
Location/Qualifiers
1. 47
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_123321.19.20.x"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 72.0%; Score 14.4; DB 28; Length 47;
Best Local Similarity 93.8%; Pred. No. 1.6e+04;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AACATGTAACCTTTG 16
40 AAACGTACTTTG 25

RESULT 2 40 bp DNA linear GSS 13-UTN-2002
BH853744 SALK_078203.44.80.x Arabidopsis thaliana TDNA insertion lines
LOCUS Arabidopsis thaliana genomic clone SALK_078203.44.80.x, genomic survey sequence.

ACCESSION BH853744.1 GI:21424615
VERSION
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)

REFERENCE
AUTHORS Ekaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 40)
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab, C., Jeske, A., Karne, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J. and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker

JOURNAL
COMMENT Salik Institute Genomic Analysis Laboratory (SIGAL)
The Salik Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salik.edu

FEATURES
source
Class: TDNA tagged.
Location/Qualifiers
1. 40
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_078203.44.80.x"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at

ORIGIN

Query Match 68.0%; Score 13.6; DB 28; Length 40;
Best Local Similarity 80.0%; Pred. No. 3.8e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AACATGTAACCTTTGTC 20
2 AAATATTAACCTTTGTC 21

RESULT 3 42 bp DNA linear GSS 24-OCT-2002
AL942024 Arabidopsis thaliana T-DNA flanking sequence GK-262601-014926,
LOCUS genomic survey sequence.
DEFINITION

ACCESSION AL942024.1 GI:24398622
VERSION
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)

REFERENCE
AUTHORS Ekaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 Strizhov, N., Li, Y., Rosso, M., Viehoveer, P., Dekker, K., Saedler, H. and Weishaar, B.
A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines
Unpublished

JOURNAL
COMMENT Rosso, M., Strizhov, N., Li, Y., Reise, B., Dekker, K. and Weishaar, B.
A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics
Unpublished
3 (bases 1 to 42)
Strizhov, N., Li, Y., Rosso, M. and Weishaar, B.
Direct Submission
Submitted (21-OCT-2002) Weishaar, B., Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence is recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At5g60000. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

source

Location/Qualifiers
1. 42
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-262601-014926"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector PAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

ORIGIN

Query Match 68.0%; Score 13.6; DB 29; Length 42;
Best Local Similarity 80.0%; Pred. No. 3.8e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AACATGTAAGTTTGGTCA 20
 |||||
 2 AACATGTAAGTTTGGTCA 21

RESULT 4
 BJ044144 35 bp mRNA linear EST 29-SEP-2003
 LOCUS BJ044144 NIBB Mochii normalized Xenopus neurula library Xenopus
 DEFINITION laevis cDNA clone X1012903 3', mRNA sequence.
 ACCESSION BJ044144
 VERSION BJ044144.1 GI:117395249
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 35)
 Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and
 Kohara, T.
 Expressed genes in X. laevis embryo
 Unpublished (2001)
 Contact: Tadao Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp
 The information of this clone is available through the following
 URL.
 http://xenopus.nibb.ac.jp.
 Location/Qualifiers
 1..35
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="X1012903"
 /cissue_type="whole embryo"
 /dev_stage="stage 15"
 /clone_1lb="NIBB Mochii normalized Xenopus neurula
 library"

ORIGIN
 Query Match 66.0%; Score 13.2; DB 12; Length 35;
 Best Local Similarity 75.0%; Pred. No. 5.7e+04;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AACATGTAAGTTTGGTCA 20
 |||||
 34 AACATGTAAGTTTGGTCA 15

RESULT 5
 A2796505 40 bp DNA linear GSS 16-FEB-2001
 LOCUS A2796505
 DEFINITION 2M0052G16F Mouse 10kb plasmid UGCGIM library Mus musculus genomic
 clone UGCG2M0052G16 F, genomic survey sequence.
 ACCESSION A2796505.1 GI:12944632
 VERSION A2796505.1
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 40)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
 Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
 Niederhausen, A. and Wright, D. Weis, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 TITLE

JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0052 row: G column: 16
 Seq primer: CGTTGTAACGACGCGCACT
 Class: plasmid ends
 High quality sequence stop: 40.
 Location/Qualifiers
 1..40
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UGCG2M0052G16"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_1lb="Mouse 10kb plasmid UGCGIM library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD2 (g1473214|g2|A129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

ORIGIN

Query Match 64.0%; Score 12.8; DB 28; Length 40;
 Best Local Similarity 87.5%; Pred. No. 9e+04;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AACATGTAAGTTTGG 17
 |||||
 22 AACATGTAAGTTTGG 37

RESULT 6
 BH911929 42 bp DNA linear GSS 04-SEP-2002
 LOCUS BH911929
 DEFINITION SRAK_073084.32.55.x Arabidopsis thaliana TDM insertion lines
 Arabidopsis thaliana genomic clone SRAK_073084.32.55.x, genomic
 survey sequence.
 ACCESSION BH911929
 VERSION BH911929.1 GI:22724862
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eusteroideae; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 42)
 Alonso, J.M., Leisner, T.J., Barajas, P., Chen, H., Cheuk, R.,
 Gadgil, C., Jeake, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L.,
 Shinn, P., Zimmerman, J. and Ecker, J.R.
 A Sequence-Indexed Library of Insertion Mutations in the
 TITLE

JOURNAL
COMMENT
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.

FEATURES
source
Class: TDNA tagged.
Location/Qualifiers

```
1..42
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_073084.32.55.x"
/clone_1lb="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna_protocols.html"
```

ORIGIN

Query Match 64.0%; Score 12.8; DB 28; Length 42;
Best Local Similarity 87.5%; Pred. No. 9e+04; 2; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AACATGTAACCTTTG 16
26 AAAATGTAACCTTTG 11

DB

RESULT 7
B0595103 46 bp mRNA linear EST 06-DEC-2002
LOCUS B0595103
DEFINITION E012708-024-023-011-SP6 MP12-ADIS-024-developing root Beta vulgaris
CDNA clone 024-023-011 5-PRIME, mRNA sequence.

ACCESSION B0595103
VERSION B0595103.1 GI:26124686
KEYWORDS EST.
SOURCE Beta vulgaris
ORGANISM Beta vulgaris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Beta.

REFERENCE 1 (bases 1 to 46)
Hervig, R., Schulz, B., Weishaar, B., Hennig, S., Steinfath, M.,
Drungowski, M., Stahl, D., Wruick, W., Menze, A., O'Brien, J., Lehnach, H.,
and Radcliff, U.

TITLE Construction of a 'unigene' cDNA clone set by oligonucleotide
fingerprinting allows access to 25 000 potential sugar beet genes
JOURNAL Plant J. 32 (5), 845-857 (2002)
MEDLINE 22362189
PUBMED 12472698

COMMENT Contact: Weishaar B
ADIS DNA core facility at MP12
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weishaar@mpiz-koeln.mpg.de
Insert Length: 46 Std Error: 0.00
Plate: 23 Row: 0 Column: 11
Seq primer: SP6; CATACGATTAGCTGACCATATAG.
Location/Qualifiers

FEATURES
source

```
1..46
/organism="Beta vulgaris"
/mol_type="mRNA"
/cultivar="KWS320 (double haploid, monogerm breeding
```

ORIGIN

Query Match 64.0%; Score 12.8; DB 13; Length 46;
Best Local Similarity 87.5%; Pred. No. 9.2e+04; 2; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AACATGTAACCTTTG 17
6 AACGGAATCTTTG 21

DB

RESULT 8
CC456918 48 bp DNA linear GSS 30-MAY-2003
LOCUS SALK_101541.52.55 x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALK_101541.52.55.x, genomic
survey sequence.

ACCESSION CC456918
VERSION CC456918.1 GI:31218116
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Brassicales; Brassicaceae; Arabidops.

REFERENCE 1 (bases 1 to 48)
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R.,
Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L.,
Shim, P., Zimmerman, J. and Ecker, J.R.

TITLE A sequence-indexed library of insertion mutations in the
Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.

FEATURES
source
Class: TDNA tagged.
Location/Qualifiers

```
1..48
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_101541.52.55.x"
/clone_1lb="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna_protocols.html"
```

ORIGIN

Query Match 62.0%; Score 12.4; DB 28; Length 48;
 Best Local Similarity 92.9%; Pred. No. 1.4e+05;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AACATGTAAGTTT 15
 |||||
 Db 12 AACATGTAAGTTT 25

RESULT 9
 AM395746 49 bp mRNA linear EST 03-DEC-2001
 LOCUS 5574h01.y1 Gm-cl007 Glycine max CDNA clone GENOME SYSTEMS CLONE ID:
 DEFINITION Gm-cl007-2186 5', mRNA sequence.
 ACCESSION AM395746
 VERSION AM395746.1 GI:6914216
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 49)
 Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Coryell, V.,
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
 Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M.,
 Bowers, Y., Peterson, B., Skaller, T., Gibbons, M., Pape, D., Harvey, N.,
 Schurk, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterson, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Putative full length read
 Vector to vector length is 50 This clone is available through:
 Reegen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL
 35801 For further information call: (800)-533-4363 or contact via
 email: cdu@reegen.com
 Seq primer: -40RP from Gibco.
 Location/Qualifiers

FEATURES

source 1..49
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl007-2186"
 /lab_host="DH10B"
 /clone_id="Gm-cl007"
 /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI. This
 CDNA library was constructed from mRNA isolated from
 immature cotyledons of greenhouse grown plants (individual
 seed fresh weight of 100-300mg). The library was prepared
 using the Life Technologies pSuperScript CDNA library
 construction kit. Complementary DNA was synthesized from
 mRNA using a poly (dT) sequence with a NotI restriction
 site. SalI linker adapters were ligated to the
 blunt-ended CDNA fragments followed by NotI digestion. The
 CDNA fragments were directionally cloned into the
 NotI-SalI restriction site of the pSPORT1 vector. The
 ligated CDNA fragments were transformed into E. coli
 Electromax DH10B host cells. This library was constructed
 by Dr. Lila Vodkin and Dr. Anu Khanna."

ORIGIN

Query Match 62.0%; Score 12.4; DB 10; Length 49;
 Best Local Similarity 92.9%; Pred. No. 1.4e+05;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AACATGTAAGTTT 14
 |||||
 Db 36 ATACATGTAAGTTT 49

RESULT 10
 TA264E04Q 23 bp DNA linear GSS 13-DEC-2000
 LOCUS T. brucei sheared genomic DNA clone 264e04, reverse sequence.
 DEFINITION
 ACCESSION TA264E04Q
 VERSION AL486592
 KEYWORDS AL486592.1 GI:11850609
 SOURCE GSS.
 ORGANISM Trypanosoma brucei
 Trypanosoma brucei
 Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma.
 1 (bases 1 to 23)
 Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,
 Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
 Melville, S.E., Rejandream, M.A. and Barrell, B.G.
 Direct Submission
 Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
 project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
 Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
 nh@sanger.ac.uk
 constructed at the Institute for Genomic Research (TIGR),
 Rockville, MD. Genomic DNA isolated from a cloned population of
 Trypanosoma brucei (TRBU927/4 GUTat 10.1) was mechanically sheared
 to give a tight size distribution (4 kb). The v + i method used for the library construction is
 described in detail in Smith, H. and Venter, J.C. (Making small
 insert libraries for whole genome shotgun sequencing projects. In
 Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
 Barrell, Oxford University Press, 1999).
 Email: nh@sanger.ac.uk
 Details of T. brucei sequencing at the Sanger Centre are available
 at http://www.sanger.ac.uk/Projects/T_brucei/.

COMMENT

TITLE
 JOURNAL

REFERENCE

1 (bases 1 to 23)
 /organism="Trypanosoma brucei"
 /mol_type="genomic DNA"
 /strain="TRBU927"
 /db_xref="taxon:5691"
 /clone="264e04"

FEATURES

source 1..23
 /organism="Trypanosoma brucei"
 /mol_type="genomic DNA"
 /strain="TRBU927"
 /db_xref="taxon:5691"
 /clone="264e04"

ORIGIN

Query Match 61.0%; Score 12.2; DB 29; Length 23;
 Best Local Similarity 82.4%; Pred. No. 1.6e+05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 ACATGTAAGTTTGGTC 19
 |||||
 Db 1 ACATGTAAGTAAGTTC 17

RESULT 11
 CC886423 40 bp DNA linear GSS 31-UTL-2003
 LOCUS SALK_148577.19.00.x Arabidopsis thaliana TDNA insertion lines
 DEFINITION Arabidopsis thaliana genomic clone SALK_148577.19.00.x, genomic
 survey sequence.
 ACCESSION CC886423
 VERSION CC886423.1 GI:33362779
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 40)

AUTHORS
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R.,
Gadriah, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L.,
Shim, P., Zimmerman, U., and Ecker, U.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within 300 bases of the 5' end of
AT2g47370.
Class: TDNA tagged.

FEATURES
Location/Qualifiers
1..40
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_148577.19.00.x"
/note="1kb=Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN
Query Match 61.0%; Score 12.2; DB 29; Length 40;
Best Local Similarity 82.4%; Pred. No. 1.7e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Db 2 AACATGTAACCTTTGGT 18
21 AAATGTAACCTATTCGT 37

RESULT 12
BX659565/c 47 bp DNA linear GSS 03-OCT-2003
LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-646G10-023086,
DEFINITION genomic survey sequence.
ACCESSION BX659565
VERSION BX659565.1 GI:37615953
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 Strizhov, N., Li, Y., Rosso, M., Viehoveer, P., Dekker, K., Saedler, H.
and Weishaar, B.
A pipeline for automated high-throughput generation of FSTs
(flanking sequence tags) from Arabidopsis thaliana T-DNA
transformed lines
Unpublished
2 Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weishaar, B.
A new Arabidopsis thaliana T-DNA tagged population (GABI-Kat)
for flanking sequence tag based reverse genetics
Unpublished
3 (bases 1 to 47)
Strizhov, N., Rosso, M., Li, Y. and Weishaar, B.
Direct Submission
Submitted (06-OCT-2003) Weishaar, B., Max-Planck-Institut fuer
Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence is recovered from the left border of the T-DNA. It

indicates an insertion within the locus defined by clone F1A17.
The sequences are generated at the MPI for Plant Breeding Research
in the context of the GABI-Kat project. GABI-Kat is part of the
German Plant Genomics program designated 'GABI'. Information on
line availability can be found at:
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>
Location/Qualifiers
1..47
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-646G10-023086"
/note="1kb=Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector pACT61. The lines contain one or more T-DNA
insertions. The DNA fragment(s) resulting from the PCR
were directly sequenced to determine the genomic sequence
flanking the insertion. Sequences displaying significant
similarity to the A. thaliana nuclear genome sequence were
processed for submission. T-DNA derived sequences were
removed"

ORIGIN
Query Match 61.0%; Score 12.2; DB 29; Length 47;
Best Local Similarity 82.4%; Pred. No. 1.8e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Db 2 AACATGTAACCTTTGGT 18
32 AATATCTAACCTTTGGT 16

RESULT 13
AZ602539 48 bp DNA linear GSS 13-DEC-2000
LOCUS 1M0421B19 Mouse 10kb plasmid UGCLM library Mus musculus genomic
DEFINITION clone UGCLM0421B19 F, genomic survey sequence.
ACCESSION AZ602539
VERSION AZ602539.1 GI:11724729
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
1 (bases 1 to 48)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Petersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von
Niederhausern, A. and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 1000 Std Error: 0.00
Place: 0421 row: B column: 19
Seq primer: CCGTGAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 48.
Location/Qualifiers
1..48
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

FEATURES
source

/clone="UDGCM0421B19"
/sex="Male"
/lab.host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone.lib="Mouse 10kb plasmid UDGCM library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114[gb]|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 61.0%; Score 12.2; DB 28; Length 48;
Best Local Similarity 82.4%; Pred. No. 1.8e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AACATGTAACCTTTGG 17
Db 25 AACATGTAACCTTTGG 41

RESULT 14 49 bp DNA linear GSS 26-NOV-2002
B2382085
LOCUS SALX_117842.35.15.x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALX_117842.35.15.x, genomic survey sequence.

ACCESSION B2382085
VERSION B2382085.1 GI:25476732
KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shim,P., Zimmerman,J., and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 49)

AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shim,P., Zimmerman,J., and Ecker,J.R.

TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome

JOURNAL Arabidopsis Genome
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (Signal)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu

FEATURES
source This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of At2g47370.
Classes: TDNA tagged.

Location/Qualifiers
1..49
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALX_117842.35.15.x"

/clone.lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contained one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 61.0%; Score 12.2; DB 28; Length 49;
Best Local Similarity 82.4%; Pred. No. 1.8e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AACATGTAACCTTTGGT 18
Db 25 AACATGTAACCTTTGGT 41

RESULT 15 49 bp DNA linear GSS 31-JUL-2003
CC886429
LOCUS SALX_148585.40.20.x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALX_148585.40.20.x, genomic survey sequence.

ACCESSION CC886429
VERSION CC886429.1 GI:33362785
KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shim,P., Zimmerman,J., and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome

Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (Signal)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu

FEATURES
source This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of At2g47370.
Classes: TDNA tagged.

Location/Qualifiers
1..49
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALX_148585.40.20.x"
/note="PCR was performed on Arabidopsis thaliana lines each of which contained one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 61.0%; Score 12.2; DB 29; Length 49;
Best Local Similarity 82.4%; Pred. No. 1.8e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AACATGTAACCTTTGGT 18
Db 25 AACATGTAACCTTTGGT 41

Mon Mar 29 09:29:17 2004

Search completed: March 27, 2004, 10:40:47
Job time : 2410 secs

us-10-006-191-48.1150.rst

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2004, 09:28:31 ; Search time 266 Seconds
(without alignments)
279.926 Million cell updates/sec

Title: US-10-006-191-48

Perfect score: 20
Sequence: 1 aaacgtgaactttgtgca 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 1421238

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PC1_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PC10_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-----------------------|
| 1 | 20 | 100.0 | 20 | 14 | US-10-006-191-48 |
| 2 | 19 | 95.0 | 20 | 14 | US-10-006-191-64 |
| 3 | 15.2 | 75.0 | 25 | 14 | US-10-098-263B-56430 |
| 4 | 15 | 75.0 | 20 | 14 | US-10-006-191-47 |
| 5 | 15 | 75.0 | 25 | 14 | US-10-098-263B-9793 |
| 6 | 14.8 | 74.0 | 25 | 14 | US-09-770-107-87 |
| 7 | 14.8 | 74.0 | 25 | 14 | US-10-098-263B-15896 |
| 8 | 14.8 | 74.0 | 27 | 15 | US-10-348-143-2469 |
| 9 | 14.2 | 71.0 | 25 | 14 | US-10-098-263B-105927 |
| 10 | 14 | 70.0 | 20 | 14 | US-10-006-191-63 |
| 11 | 13.6 | 68.0 | 25 | 14 | US-10-098-263B-17802 |
| 12 | 13.6 | 68.0 | 25 | 14 | US-10-098-263B-56429 |
| 13 | 13.6 | 68.0 | 25 | 14 | US-10-098-263B-57645 |
| 14 | 13.6 | 68.0 | 35 | 10 | US-09-309-196-15 |
| 15 | 13.6 | 68.0 | 35 | 10 | US-09-201-396-10 |

| Query | Match | Score | DB ID | Length | Sequence |
|-------|--------|-------|-------|--------|-----------------------|
| 1 | 100.0% | 20 | 14 | 20 | US-10-006-191-48 |
| 2 | 95.0% | 19 | 14 | 20 | US-10-006-191-64 |
| 3 | 75.0% | 15.2 | 14 | 25 | US-10-098-263B-56430 |
| 4 | 75.0% | 15 | 14 | 20 | US-10-006-191-47 |
| 5 | 75.0% | 15 | 14 | 25 | US-10-098-263B-9793 |
| 6 | 74.0% | 14.8 | 14 | 25 | US-09-770-107-87 |
| 7 | 74.0% | 14.8 | 14 | 25 | US-10-098-263B-15896 |
| 8 | 74.0% | 14.8 | 15 | 27 | US-10-348-143-2469 |
| 9 | 71.0% | 14.2 | 14 | 25 | US-10-098-263B-105927 |
| 10 | 70.0% | 14 | 14 | 20 | US-10-006-191-63 |
| 11 | 68.0% | 13.6 | 14 | 25 | US-10-098-263B-17802 |
| 12 | 68.0% | 13.6 | 14 | 25 | US-10-098-263B-56429 |
| 13 | 68.0% | 13.6 | 14 | 25 | US-10-098-263B-57645 |
| 14 | 68.0% | 13.6 | 10 | 35 | US-09-309-196-15 |
| 15 | 68.0% | 13.6 | 10 | 35 | US-09-201-396-10 |

ALIGNMENTS

| Query | Match | Score | DB ID | Length | Sequence |
|-------|--------|-------|-------|--------|-----------------------|
| 1 | 100.0% | 20 | 14 | 20 | US-10-006-191-48 |
| 2 | 95.0% | 19 | 14 | 20 | US-10-006-191-64 |
| 3 | 75.0% | 15.2 | 14 | 25 | US-10-098-263B-56430 |
| 4 | 75.0% | 15 | 14 | 20 | US-10-006-191-47 |
| 5 | 75.0% | 15 | 14 | 25 | US-10-098-263B-9793 |
| 6 | 74.0% | 14.8 | 14 | 25 | US-09-770-107-87 |
| 7 | 74.0% | 14.8 | 14 | 25 | US-10-098-263B-15896 |
| 8 | 74.0% | 14.8 | 15 | 27 | US-10-348-143-2469 |
| 9 | 71.0% | 14.2 | 14 | 25 | US-10-098-263B-105927 |
| 10 | 70.0% | 14 | 14 | 20 | US-10-006-191-63 |
| 11 | 68.0% | 13.6 | 14 | 25 | US-10-098-263B-17802 |
| 12 | 68.0% | 13.6 | 14 | 25 | US-10-098-263B-56429 |
| 13 | 68.0% | 13.6 | 14 | 25 | US-10-098-263B-57645 |
| 14 | 68.0% | 13.6 | 10 | 35 | US-09-309-196-15 |
| 15 | 68.0% | 13.6 | 10 | 35 | US-09-201-396-10 |

CURRENT APPLICATION NUMBER: US/10/006,191
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 153
SEQ ID NO 64
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
US-10-006-191-64

Query Match 95.0%; Score 19; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACATGTAACCTTTGGTC 19
DB 2 AAACATGTAACCTTTGGTC 20

RESULT 3

US-10-098-263B-56430
Sequence 56430, Application US/10098263B
Publication No. US20030104410A1
GENERAL INFORMATION:
APPLICANT: Miltman, Michael
TITLE OF INVENTION: Human Microarray
FILE REFERENCE: 3118.1
CURRENT APPLICATION NUMBER: US/10/098,263B
CURRENT FILING DATE: 2003-01-08
PRIOR APPLICATION NUMBER: 60/276,759
PRIOR FILING DATE: 2001-03-16
NUMBER OF SEQ ID NOS: 13,066
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 56430
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapien
US-10-098-263B-56430

Query Match 76.0%; Score 15.2; DB 14; Length 25;
Best Local Similarity 85.0%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAACATGTAACCTTTGGTC 20
DB 1 AGACATGTAACCTTTTTC 20

RESULT 4

US-10-006-191-47
Sequence 47, Application US/10006191
Publication No. US20030144223A1
GENERAL INFORMATION:
APPLICANT: William Gaarde
TITLE OF INVENTION: ANTISENSE MODULATION OF CONNECTIVE TISSUE GROWTH FACTOR EXPRESSION
FILE REFERENCE: RTS-0274
CURRENT APPLICATION NUMBER: US/10/006,191
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 153
SEQ ID NO 47
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
US-10-006-191-47

Query Match 75.0%; Score 15; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TGTACTTTGGTC 20
DB 1 TGTACTTTGGTC 15

RESULT 5

US-10-098-263B-49793/c
Sequence 49793, Application US/10098263B
Publication No. US20030104410A1
GENERAL INFORMATION:
APPLICANT: Miltman, Michael
TITLE OF INVENTION: Human Microarray
FILE REFERENCE: 3118.1
CURRENT APPLICATION NUMBER: US/10/098,263B
CURRENT FILING DATE: 2003-01-08
PRIOR APPLICATION NUMBER: 60/276,759
PRIOR FILING DATE: 2001-03-16
NUMBER OF SEQ ID NOS: 13,066
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 49793
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapien
US-10-098-263B-49793

Query Match 75.0%; Score 15; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ACATGTAACCTTTGG 17
DB 15 ACATGTAACCTTTGG 1

RESULT 6

US-09-770-107-87/c
Sequence 87, Application US/09770107
Publication No. US20030054345A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Weyer, Joanne
APPLICANT: Barrington-Martin, Rory
APPLICANT: Parker, Alexander
APPLICANT: Barnes, Glenn
TITLE OF INVENTION: Compositions and methods for the diagnosis and treatment of neuropsychiatric disorders, including schizophrenia
FILE REFERENCE: 3322/05401
CURRENT APPLICATION NUMBER: US/09/770,107
CURRENT FILING DATE: 2001-01-24
NUMBER OF SEQ ID NOS: 127
SOFTWARE: PatentIn version 3.0
SEQ ID NO 87
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapiens
US-09-770-107-87

Query Match 74.0%; Score 14.8; DB 10; Length 25;
Best Local Similarity 88.9%; Pred. No. 2e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAACATGTAACCTTTGGT 18
DB 21 AAACATGTAACGTTGGT 4

RESULT 7

US-10-098-263B-15896
Sequence 15896, Application US/10098263B
Publication No. US20030104410A1
GENERAL INFORMATION:
APPLICANT: Miltman, Michael
TITLE OF INVENTION: Human Microarray

```
FILE REFERENCE: 3118.1
CURRENT APPLICATION NUMBER: US/10/098,263B
CURRENT FILING DATE: 2003-01-08
PRIOR APPLICATION NUMBER: 60/276,759
PRIOR FILING DATE: 2001-03-16
NUMBER OF SEQ ID NOS: 131066
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO: 15896
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapien
US-10-098-263B-15896

Query Match
Best Local Similarity 88.9%; Pred. No. 2e+03; Length 25;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AACATGTAACCTTTGGTCA 20
Db 4 ACATGTTACTTTGTTCA 21

RESULT 8
US-10-349-143-2469/c
Sequence 2469, Application US/10349143
Publication No. US20040005584A1
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
FILE REFERENCE: GENSET.020CPI
CURRENT FILING DATE: US/10/349,143
CURRENT APPLICATION NUMBER: 2003-01-21
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: US/09/422,978
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: US 09/238,650
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: US 60/109,732
PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
NUMBER OF SEQ ID NOS: 11796
SEQ ID NO: 2469
LENGTH: 47
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 24
OTHER INFORMATION: 99-11179-239 : polymorphic base C or T
US-10-349-143-2469

Query Match
Best Local Similarity 74.0%; Score 14.8; DB 15; Length 47;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AACATGTAACCTTTGGTCA 20
Db 33 AACATATACCTTTGTGA 14

RESULT 9
US-10-098-263B-105927
Sequence 105927, Application US/10098263B
Publication No. US20030104410A1
GENERAL INFORMATION:
APPLICANT: Miltman, Michael
TITLE OF INVENTION: Human Microarray
FILE REFERENCE: 3118.1
CURRENT APPLICATION NUMBER: US/10/098,263B
CURRENT FILING DATE: 2003-01-08
PRIOR APPLICATION NUMBER: 60/276,759

Query Match
Best Local Similarity 80.0%; Pred. No. 2.2e+03; Length 25;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AACATGTAACCTTTGGTCA 20
Db 6 AAAAGTAACCTTTGGTCA 24

RESULT 10
US-10-006-191-63
Sequence 63, Application US/10006191
Publication No. US20030144223A1
GENERAL INFORMATION:
APPLICANT: William Garde
APPLICANT: Andrew T. Watt
TITLE OF INVENTION: ANTISENSE MODULATION OF CONNECTIVE TISSUE GROWTH FACTOR EXPRESSIO
FILE REFERENCE: RTS-0274
CURRENT APPLICATION NUMBER: US/10/006,191
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 153
SEQ ID NO: 63
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
US-10-006-191-63

Query Match
Best Local Similarity 70.0%; Score 14; DB 14; Length 20;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 GTAACCTTTGGTCA 20
Db 1 GTAACCTTTGGTCA 14

RESULT 11
US-10-098-263B-17802/c
Sequence 17802, Application US/10098263B
Publication No. US20030104410A1
GENERAL INFORMATION:
APPLICANT: Miltman, Michael
TITLE OF INVENTION: Human Microarray
FILE REFERENCE: 3118.1
CURRENT APPLICATION NUMBER: US/10/098,263B
CURRENT FILING DATE: 2003-01-08
PRIOR APPLICATION NUMBER: 60/276,759
PRIOR FILING DATE: 2001-03-16
NUMBER OF SEQ ID NOS: 131066
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO: 17802
LENGTH: 25
TYPE: DNA
ORGANISM: Homo Sapien
US-10-098-263B-17802

Query Match
Best Local Similarity 68.0%; Score 13.6; DB 14; Length 25;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AACATGTAACCTTTGGTCA 20
```


Db 20 AAACATCTAAGCTTGTGTA 1

RESULT 12

US-10-098-263B-56429
; Sequence 56429, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Miltman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 56429
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-56429

Query Match

68.0%; Score 13.6; DB 14; Length 25;
Best Local Similarity 80.0%; Pred. No. 7.3e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAACATCTAAGCTTGTGTA 20
Db 1 AGACATGTAAGCTTGTGTA 20

RESULT 13

US-10-098-57645/C
; Sequence 57645, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Miltman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 57645
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-57645

Query Match

68.0%; Score 13.6; DB 14; Length 25;
Best Local Similarity 80.0%; Pred. No. 7.3e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAACATGTAAGCTTGTGTA 20
Db 23 AAACATGTAAGCTTGTGTA 4

RESULT 14

US-09-309-196-15/C
; Sequence 15, Application US/09309196
; Publication No. US20030008380A1
; GENERAL INFORMATION:
; APPLICANT: FOWLES, Dana M.
; APPLICANT: BROACH, Jim
; APPLICANT: MANFREDI, John
; APPLICANT: KLEIN, Christine
; APPLICANT: MURPHY, Andrew J.
; APPLICANT: PAUL, Jeremy
; APPLICANT: TRUEHEART, Joshua

TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE
; PHEROMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR

NUMBER OF SEQUENCES: 119

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK

STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/309,196

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/322,137

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/190,328

FILING DATE: 31-JAN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/041,431

FILING DATE: 31-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: COOPER, Iver P.

REGISTRATION NUMBER: 28,005

REFERENCE/DOCKET NUMBER: FOLWES-2C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

TELEX: 248633

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 35 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: synthetic DNA

US-09-309-196-15

Query Match

68.0%; Score 13.6; DB 10; Length 35;
Best Local Similarity 80.0%; Pred. No. 7.7e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAACATGTAAGCTTGTGTA 20
Db 20 AAACATGTAAGCTTGTGTA 1

RESULT 15

US-09-201-396-10/C
; Sequence 10, Application US/09201396A
; Publication No. US20030009022A1
; GENERAL INFORMATION:
; APPLICANT: Klein, Christine A.
; APPLICANT: Murphy, Andrew J.
; APPLICANT: Paul, Jeremy
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING RECEPTOR
; FILE REFERENCE: CPT-012C99
; CURRENT APPLICATION NUMBER: US/09/201,396A
; CURRENT FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: 08/582,333
; EARLIER FILING DATE: 1996-01-17
; EARLIER APPLICATION NUMBER: 08/322,137
; EARLIER FILING DATE: 1994-10-13
; EARLIER APPLICATION NUMBER: 08/309,313
; EARLIER FILING DATE: 1994-09-20

; EARLIER APPLICATION NUMBER: 08/190,328
 ; EARLIER FILING DATE: 1994-01-31
 ; EARLIER APPLICATION NUMBER: 08/041,431
 ; EARLIER FILING DATE: 1993-03-31
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO: 10
 ; LENGTH: 35
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: Construct
 US-09-201-396-10

Query Match 68.0%; Score 13.6; DB 10; Length 35;
 Best local Similarity 80.0%; Pred. No. 7; +03;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAACATGTAACCTTTGGTCA 20
 |||||
 DB 20 AAACATGTAACCTTTGGTCA 1

Search completed: March 27, 2004, 10:46:54
 Job time : 267 secs